Representation Schemes for Investigating Non-Linear Processes

by
DONALD B. DeVORKIN

MASSACHUSETTS INSTITUTE OF TECHNOLOGY
DEPARTMENT OF METEOROLOGY

Contract No. AF19(604)-4969
Project No. 8628
Task No. 86281

Scientific Report No. 1
STATISTICAL FORECASTING PROJECT
DECEMBER 1962

Prepared
for
GEOPHYSICS RESEARCH DIRECTORATE
AIR FORCE CAMBRIDGE RESEARCH LABORATORIES
OFFICE OF AEROSPACE RESEARCH
UNITED STATES AIR FORCE
BEDFORD, MASSACHUSETTS

Representation Schemes for Investigating Non-Linear Processes

by

DONALD B. DeVORKIN

MASSACHUSETTS INSTITUTE OF TECHNOLOGY
DEPARTMENT OF METEOROLOGY

Contract No.	AF19(604)-4969
Project No.	8628
Task No.	86281

Scientific Report No. 1

STATISTICAL FORECASTING PROJECT

DECEMBER 1962

Prepared
for
GEOPHYSICS RESEARCH DIRECTORATE
AIR FORCE CAMBRIDGE RESEARCH LABORATORIES
OFFICE OF AEROSPACE RESEARCH
UNITED STATES AIR FORCE
BEDFORD, MASSACHUSETTS

Representation Schemes for Investigating Non-linear Processes

by

Donald DeVorkin

ABSTRACT

The work is concerned with systems of ordinary differential equations. A framework is developed in which a statistical approach to the analysis of the equations is natural. One seeks the correlations of the solution in terms of the correlations of initial conditions and/or forcing functions. General representation schemes are developed for this purpose. It is shown that general schemes can converge for arbitrarily large but finite time. In some special cases, all-time representations are obtained.

TABLE OF CONTENTS

INTRODUCTION

(0.1) Preliminary Discussion

At the present time, we are able to enjoy the results of extensive successful application of linear methods to science and engineering. This is not to say that even essentially linear problems necessarily yield to the developed analytical methods of mathematical physics. One example of what is meant is that we can solve Laplace's equation analytically in only very special coordinate systems. Nevertheless, since we can solve many linear problems, we are sometimes able to deduce general properties for a class of problems, suggest meaningful experiments, and obtain successful engineering methodology.

The situation for non-linear problems is not nearly so satisfactory. Some very special problems have been thoroughly treated, but for the most part, analytical methods have been available to English readers only since our introduction to the Russian School by Minorsky (1947). Since then, although numerous books have appeared, the analytical methods available to us at the present are found to be substantially those in Minorsky (1947). These methods have one common aspect: The equations must, right to begin with, be in some sense nearly linear in order that the methods apply.

In this thesis, we will attack certain classes of non-linear problems from a different point of view. While one of the results is of a quasi-linear nature, the rest are truly non-linear in character. These results

contain their own inherent limitations, and hence will channel the kinds of physical problems to be considered.

(0.2) The Statistical Viewpoint

Wiener (1958) developed a means for synthesizing a class of non-linear networks. We can characterize this class by the properties:

a) Bounded inputs give bounded outputs.

b) At the present time, the output does not depend on the nature of the input acting at the infinite past.

For linear networks, one has become used to the impulse as a probe. That is, we often characterize linear networks by their impulse response. Wiener shows that for his class of non-linear problems, Brownian motion is a suitable probe, and that one can synthesize these systems when in possession of the Brownian motion response. One describes Brownian motion by its statistical properties and what is really required for the synthesis procedure are input-output cross correlations.

Let us try this sort of thing for a linear system

$$y(t) = \int_{-\infty}^{\infty} h(t-\tau)\, x(\tau)\, d\tau$$

where $x(t)$ is any representative member of an ergodic ensemble,

and $\int_{-\infty}^{\infty} |h(t)|\, dt$ exists and $< \infty$.

Then

$$\overline{y(t)\,x(t+\sigma)} = \int_{-\infty}^{\infty} h(t-\tau)\,\overline{x(\tau)\,x(t+\sigma)}\,d\tau$$

The bar denotes ensemble = time average.

For the case where the ensemble is white noise

$$\overline{x(t+\sigma)\,x(\tau)} = \delta(\tau,\, t+\sigma)$$

$$\overline{y(t)\,x(t+\tau)} = \int_{-\infty}^{\infty} h(t-\tau)\,\delta(\tau, t+\sigma)\,d\tau$$

$$= h(-\sigma)$$

To obtain $h(t)$, any ergodic ensemble of functions whose auto-correlation is a delta-function would suffice. Since the average $\overline{y(t)\,x(t+\sigma)}$ can be taken as well over time as down the ensemble, only one representative member of the ensemble need be required.

For the linear system, the higher order moments of the input can be left arbitrary. The Wiener theory for non-linear synthesis suggests Brownian motion, all correlations must be specified and there is no arbitrariness at all. Schetzen (1962) has shown that if a non-linear network uses up to n-th order products of the input, then a suitable probe is a representative member of an ergodic ensemble whose first

2n correlations are the same as for Brownian motion, but whose higher order correlations are arbitrary.

Through the above ideas, which are primarily concerned with network synthesis, we can get a look at a fundamental issue in analysis. To do this we need only to add something about the nature of a Brownian motion ensemble. Each member of a Brownian motion ensemble is continuous and defined over all time. Now let $f(t)$ be a continuous function defined on an interval of length L. Then almost every member of the ensemble approximates $f(t)$ arbitrarily closely on some interval of length L In a sense, a Brownian motion ensemble is very large. To delineate what is meant, we construct a very simple ensemble.

Let there correspond to the α th member of a Brownian motion ensemble, a function $f_\alpha(t) = C$, a constant independent of α. Our new ensemble has as many members as the Brownian motion ensemble, but displays no variety of form. The greater the variety, the more an ensemble will resemble Brownian motion. Then from above, as a network employes higher and higher order non-linearity, the variety of form displayed by an input must become greater when it is required that the response of the network to the input be capable of characterizing the behaviour of the network.

We are led to the idea that for a system which makes use of all orders of non-linearity in the input, every input-output situation must be analysed as a special case when deterministic statements are required. Hence, we shall in this work, attempt analysis ammenable to statistical

description of the output or solution in terms of statistics of the input.

(0.3) Differential Equations

The non-linear systems which shall concern us here will be supposed given by a differential equation or by a system of differential equations. Accordingly, in this section we mention a few standard ideas of use in dealing with ordinary differential equations. To begin, consider the linear system of n equations

$$\frac{dx}{dt} = A(t)x + f(t)$$

where x and f are column n-vectors and A is an nxn matrix whose elements may be functions of t, and $\frac{dx}{dt}$ is a column vector whose components are $\frac{dx_i}{dt}$

Let B be any nxn matrix. We define

$$e^B = I + \frac{B}{1!} + \frac{B^2}{2!} + \frac{B^3}{3!} + \cdots$$

where I is the identity matrix. If the elements of B are all finite, then the series which defines each element of the matrix e^B converges absolutely.

Note that in general $e^A e^B \neq e^B e^A$ unless A and B

commute. For the case $AB = BA$

$$e^A e^B = e^B e^A = e^{(A+B)}$$

These statements are deducible directly from the series definition of e^A. Since A commutes with $(-A)$,

$$e^{-A} e^A = e^{(A-A)} = I$$

It will be a help later on to have at our disposal, the Jordan Cannonical Form for matrices (Friedman, 1956). However, since the statement of the theorem is rather involved, we will here state a simple case.

Let A be any nxn matrix whose n eigenvalues are all distinct. Then there exists the decomposition of A

$$A = E D E^{-1}$$

where D is a matrix all of whose elements off the principal diagonal are zero, while the elements of the principal diagonal are the n distinct eigenvalues of A. The columns of E are the right eigenvectors of A. The rows of E^{-1} turn out to be the left eigenvectors of A.

Let t be a scalar, and A a constant matrix. Then

$$e^{At} = I + \frac{At}{1!} + \frac{A^2t^2}{2!} + \cdots$$

where At is a matrix each of whose components are those of A multiplied by t .

One can define a derivative as follows:

$$\frac{d}{dt} e^{At} = \frac{d}{dt}\left[I + \frac{At}{1!} + \frac{A^2t^2}{2!} + \cdots\right]$$

$$= \left[A + \frac{A^2t}{1!} + \frac{A^3t^2}{2!} + \cdots\right]$$

$$= A\, e^{At}$$

The Jordan Cannonical Form allows a simplification of representation for the matrix e^A .

If $A = EDE^{-1}$

$$e^A = e^{EDE^{-1}}$$

$$= I + \frac{EDE^{-1}}{1!} + \frac{(EDE^{-1})^2}{2!} + \cdots$$

$$= EIE^{-1} + \frac{EDE^{-1}}{1!} + \frac{ED^2E^{-1}}{2!} + \frac{ED^3E^{-1}}{3!} + \cdots$$

$$= E\left[I + \frac{D}{1!} + \frac{D^2}{2!} + \frac{D^3}{3!} + \cdots\right]E^{-1}$$

$$= E\, e^D E^{-1} \quad .$$

Since the eigenvalues of A are assumed all distinct,

$$e^{D} = \begin{pmatrix} e^{\lambda_1} & 0 & 0 & \cdot & \cdot & 0 \\ 0 & e^{\lambda_2} & 0 & \cdot & \cdot & 0 \\ 0 & 0 & e^{\lambda_3} & \cdot & \cdot & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & 0 & \cdot & \cdot & e^{\lambda_n} \end{pmatrix}$$

This is a simple and convenient form of e^{A}. It particularly is useful in handling ordinary differential equation systems which we now take up.

First, consider the single n-th order equation

$$\frac{d^n x}{dt^n} = f\left[\frac{d^{n-1}x}{dt^{n-1}}, \ldots, \frac{dx}{dt}, x\right]$$

The function on the right can of course be non-linear. Let

$$x = x_1$$

$$\frac{dx}{dt} = x_2 = \frac{dx_1}{dt}$$

$$\cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot$$

$$\frac{d^{n-1}x}{dt^{n-1}} = x_n = \frac{dx_{n-1}}{dt}$$

We now have a system of equations

$$\frac{dx_n}{dt} = f(x_1, x_2, \ldots, x_n)$$

$$\frac{dx_{n-1}}{dt} = x_n$$

$$\cdot \quad \cdot \quad \cdot \quad \cdot$$

$$\frac{dx_1}{dt} = x_2$$

By the change of variables introduced, the single n-th order equation is transformed into a system of n first order differential equations. However, it is not always possible to transform the first order system

$$\frac{dx_1}{dt} = f_1(x_1, \ldots, x_n)$$

$$\cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot$$

$$\frac{dx_n}{dt} = f_n(x_1, \ldots, x_n)$$

into a single n-th order equation of the form

$$\frac{d^n x}{dt^n} = f\left(\frac{d^{n-1}x}{dt^{n-1}}, \ldots, \frac{dx}{dt}, x\right)$$

For this reason, it will be preferable to work with systems of first order equations.

Let A be an nxn matrix whose components are constants independent of time. The linear homogeneous system

$$\frac{dx}{dt} = Ax, \qquad x(0) = x_0$$

is solvable as

$$x(t) = e^{At}x_0 \quad \text{since from above} \quad \frac{d}{dt}(e^{At}x_0) =$$

$$= Ae^{At}x_0 = Ax \quad \text{while} \quad e^{A(0)} = I$$

so that $x(0) = Ix_0 = x_0$, the arbitrary initial condition vector.

For A still a constant matrix and $f(t)$ a vector, a particular solution to the inhomogeneous linear system

$$\frac{dx}{dt} = Ax + f(t)$$

is $$x_p(t) = \int_0^t e^{A(t-\tau)} f(\tau)\, d\tau$$

where the integration is carried out on a term by term basis of the n-vector, $e^{A(t-\tau)}f(\tau)$. Note that $x_p(0) = 0$. That $x_p(t)$ is indeed a particular solution can be calculated directly.

The complete solution which reduces to an arbitrary vector x_o at $t = 0$ is thus

$$x(t) = e^{At}x_o + \int_0^t e^{A(t-\tau)}f(\tau)\,d\tau$$

Consider next the system of first order non-linear differential equations

$$\frac{dx}{dt} = Ax + f(x,t)$$

where A is again a constant matrix and $f(x,t)$ is a vector whose components are functions of t and the components of x. It is also supposed that $x(0) = x_o$. Then from above

$$x(t) = e^{At}x_o + \int_0^t e^{A(t-\tau)}f[x(\tau),\tau]\,d\tau$$

It will be convenient later to use this equivalent integral equation formulation. But so far we have dealt only with systems whose linear part has constant coefficients.

For the linear system of n equations

$$\frac{dx}{dt} = A(t)x$$

it is known from the theory of ordinary differential equations that over any finite time interval, there exists n independent solutions which can be arranged as the columns of the nxn matrix $H(t)$, and that

any solution can be expressed as

$$x(t) = H(t)\, y_o$$

Further, the inverse of $H(t)$ exists for all t in the interval.

$$H(0)^{-1}\, x(0) = y_o$$

$$x(t) = H(t)\, H^{-1}(0)\, x(0)$$

In an analogous manner, the system

$$\frac{dx}{dt} = A(t) + f(x, t)$$

is transformed into the integral equation

$$x(t) = H(t)\, H^{-1}(0)\, x(0) + H(t) \int_0^t H^{-1}(\tau)\, f[x(\tau), \tau]\, d\tau$$

When $f(x, t)$ does not depend on x, the integral equation becomes that particular solution to the linear inhomogeneous system

$$\frac{dx}{dt} = A(t)\, x + f(t)$$

which reduces to $x(0)$ at $t = 0$.

(0.4) All-Time Representations and Volterra Functionals

A functional is a mapping of a function space into the real line. Let us take the function space to be the set X of continuous functions $f(t)$ with $[a \leq t \leq b]$. Then to every $f(t)$, is made to correspond a real number $F[f(t)]$, called a functional.

We call an expression of the form $\int_a^b K(t) f(t)\, dt$ a regular linear functional on the function space X. $K(t)$ is fixed and assumed continuous. A regular homogeneous functional of n^{th} degree is defined analogously as:

$$\int_a^b \cdots \int_a^b K(t_1, \cdots, t_n) f(t_1) \cdots f(t_n)\, dt_1 \cdots dt_n$$

Let $F_n[f]$ be a regular homogeneous functional of n^{th} degree. Then the form $\sum_0^N F_n[f]$ is called a functional polynomial. For the usual familiar polynomials, we have the following important representation (or approximation) theorem of Weierstrass as stated by Apostol (1957).

Let f be real-valued and continuous on a closed interval $[a, b]$. Then given any $\epsilon > 0$, there exists a polynomial

(which may depend on ϵ) such that

$$|f(t) - P(t)| < \epsilon \quad \text{for every } t \text{ in } [a, b] .$$

The extension of this theorem to functionals was already known in 1910 by Frechet, and in all probability only shortly after that by Volterra. We will state the functional polynomial approximation theorem, but must first discuss continuity for functionals.

Let X be the function space defined above. Further, suppose a metric defined on X ; the distance between $f(t)$ and $g(t)$ is denoted by $\|f - g\|$. Then a functional F on X is continuous at f if for any $\epsilon > 0$, there exists a $\delta(f, \epsilon)$ such that

$$|F[f] - F[g]| < \epsilon \quad \text{whenever} \quad \|f - g\| < \delta .$$

The extension of the Weierstrass theorem to functionals is stated in Volterra (1959). However, the statement there is loose, and we will not directly quote him.

Let the function space X be as defined above; and, let there be a metric on X such that any homogeneous regular functional on X is continuous. Further, suppose that with respect to the metrix on X , $F[f]$ is <u>any</u> continuous functional, and that C is a compact set in X . Then given any $\epsilon > 0$, there exists a functional

polynomial $P[f]$ (which may depend on ϵ) such that

$$|F[f] - P[f]| < \epsilon \quad \text{for all } f \text{ in } C .$$

It is seen that functional polynomials possess the same representation qualities for functionals as do polynomials for ordinary functions.

The Stone-Weierstrass theorem is an abstract general statement which contains the above theorem as a special case. Appendix C makes use of the Stone-Weierstrass theorem. For the time being, let it be sufficient to say that Brilliant's (1958) topology, used in conjunction with the Stone-Weierstrass theorem, indicates that there may be a class of functionals, occurring commonly in practice, which can be represented by functional power series of the form $\sum_{n=0}^{\infty} P_n[f]$

where

$$P_n[f] = \int_{t_1=-\infty}^{0} \cdots \int_{t_n=-\infty}^{0} K[t_1, \cdots, t_n] \cdot f(t_1) \cdots f(t_n)\, dt_1 \cdots dt_n$$

The difference here is that the functions $f(t)$ of the function space X are now defined over the infinite interval $(-\infty < t \leq 0]$

As mentioned earlier, Wiener (1958) has given a procedure for obtaining the kernels in the integrals of a related scheme when such a representation

exists. There are certain fundamental unresolved difficulties – even when the representation exists – which may be somewhat relieved with the tools developable with the insertion of Brilliant's (1958) topology into the Stone-Weierstrass theorem. We will pursue these lines further in the third chapter. Although many useful non-linear operators may be representable by a Wiener type scheme, we shall here deal with problems for which nothing so nice exists as a uniformly convergent over X functional power series representation when $f(t)$ in X is defined over $(-\infty < t \leq 0]$. Our interest lies in non-linear ordinary differential equations of the kind

$$\frac{dx}{dt} = f(x,t)$$

where x is an n-vector, and f is an n-vector whose components are in general non-linear functions of t and the components of x .

For simplicity, let the equation take the form

$$\frac{dx}{dt} = f(x) + g(t)$$

where $g(t)$ will be some prescribed function of time and x, f, and g are one-vectors. Now in addition, let the equation

$$\frac{dx}{dt} = f(x)$$

possess two asymptotically stable limit points, x_1 and x_2 (x_1 and x_2 are one-vectors). Associated with x_1 is a region R_1 and with x_2 a region R_2 such that if $x(t)$ is in R_1 for some T, then $x(t)$ stays in R_1 for all $t > T$ and asymptotically $x(t) \to x_1$; the same situation obtains of course for R_2. The two regions are clearly disjoint.

We suppose the process to have been going on from minus infinity until now, and that there is a functional power series representation for t = now = 0. The initial conditions at $-\infty$ are supposed fixed independent of the forcing function $g(t)$. If $\frac{dx}{dt} = F(x) + g(t)$, then

$$x(0) = K_0 + \int_{-\infty}^{0} K_1(t)\, g(t)\, dt + \int_{-\infty}^{0}\int_{-\infty}^{0} K_2(t_1, t_2)\, g(t_1)\, g(t_2)\, dt_1\, dt_2 + \cdots$$

where K_0 is a constant. This is the functional power series. Also if $\frac{dy}{dt} = F(y) + h(t)$, then we must be able to write

$$y(0) = K_0 + \int_{-\infty}^{0} K_1(t)\, h(t)\, dt + \int_{-\infty}^{0}\int_{-\infty}^{0} K_2(t_1, t_2)\, h(t_1)\, h(t_2)\, dt_1\, dt_2 + \cdots$$

$$= \sum_{0}^{\infty} K_n[h]$$ in an obvious notation.

There exists an ϵ such that an ϵ – neighborhood about the asymptotically stable limit point x_1 is contained in R_1, and an ϵ-neighborhood about x_2 is contained in R_2. Also, there exists an $N(\epsilon)$ such that

$$\left| x(0,f) - \sum_{0}^{M} K_n[f] \right| < \epsilon \qquad \text{for all}$$

$M > N$ and all f of the function space.

We now choose special forcing functions. Let $g(t)$ drive the solution $x(t)$ into region R_1 at time T and then turn off; i.e., $g(t) = 0$ for $t > T$. Similarly let $h(t)$ drive $x(t)$ into R_2 at $t = T$ and then turn off. It is necessary here to take $g(t)$ and $h(t)$ bounded so that by having T far enough in the past and requiring that all individual integrals exist we can have

$$|x_1 - x(0)| < \frac{\epsilon}{3}$$

$$\left| x(0) - \sum_{n=0}^{N} K_n[g] \right| < \frac{\epsilon}{3}$$

$$\left| \sum_{n=1}^{N} K_n[g] \right| < \frac{\epsilon}{3}$$

$$\left| x_2 - y(0) \right| < \frac{\epsilon}{3}$$

$$\left| y(0) - \sum_{m=0}^{N} K_m[h] \right| < \frac{\epsilon}{3}$$

$$\left| \sum_{m=1}^{N} K_m[h] \right| < \frac{\epsilon}{3}$$

This implies that K_0 is within ϵ of both x_1 and x_2 which is the required contradiction. It is to be noted that in any individual problem, we may not be able to choose parameters, forcing functions, etc. as required above. However, there will be <u>some</u> not at all pathological equations for which we can proceed in the above fashion. This is all that is needed to invalidate functional power series expansions as a general approach. Yet, for certain statistical investigations, one is reluctant indeed to give up functional power series schemes; for the forcing function appears in products with itself and drawing averaging bars yields output correlations in terms of input correlations directly. Hence in the remainder of the paper, we shall pursue certain areas where the special, but nevertheless important, situations allow of a representation in a form particularly suited to statistical investigation.

(0.5) Content and Form of the Thesis

The introduction itself is to serve two main purposes. The first of these is to introduce the reader to the problems and limitations

involved in attempting to obtain general solutions of non-linear equations. These difficulties have caused the author to adopt certain points of view. For example, the discussion on the statistical approach in section (0.2). This, and certain ideas stressed in the introduction, should serve to orient the reader's point of view sufficiently to obviate the need for re-stressing in the remainder of the paper. So that again by way of example, it will be supposed throughout the body of the thesis that the reader is himself capable of drawing ensemble average bars in order to obtain statistics as any given situation does or does not require.

The second major purpose of the introduction is mathematical in nature. Here it is hoped that a reader not already adequately prepared will gain some mathematical ideas and manipulations to serve as background for the thesis. Much of this material perhaps need not be explicitly used in reading through the body of the thesis. Nevertheless, it may prove valuable background, especially should the reader wish to use the results on his own problems.

Chapter I deals with a pertubation procedure. The method is inherently quasi-linear in nature. However, it does give an important area for which an all-time representation is possible. Chapters II and III give results which are truly non-linear in character. The general statements will here hold for only an arbitrarily large but still finite length of time. Of course from section (0.4) this is to be expected.

In Chapter II, a condition is given, the satisfaction of which allows of approximate solutions which converge uniformly over all time.

This restriction is simply that for any finite initial conditions, and zero forcing function, the system damps asymptotically to zero. Under this condition, Chapter II gives an all time representation scheme for analytic systems without forcing. When the condition is not fulfilled, the scheme converges uniformly over any arbitrarily large but finite time interval.

Chapter III presents a finite time representation scheme for analytic systems with forcing. The difficulties of an all-time Volterra expansion are again considered.

The physical problems discussed have been included only to illustrate applicability of the abstract statements. In each case, the problem is briefly set up. It is hoped that later work will carry out the indicated procedures and obtain physically significant results.

One further remark about the examples. Field problems of the kind dealt with here are formulated by partial differential equations. These represent a non-denumerable infinity of equations. However, we shall always suppose that high frequency effects are negligible and that for physical purposes we could approximately represent the problem by a finite dimensional mesh space, or by say a finite number of Fourier coefficients. The reader can suppose that any infinite space is approximable by a finite one, so that Fourier transforms of probabilistic functions necessarily exist. However, the theory of formal Fourier representations (Lighthill, 1959) can be invoked should an infinite space and rigorous approach be desired.

Finally, proofs which are unduly long, complicated, or mathemathical have been relegated to the Appendices. The Appendices themselves are mathematical in nature. Terms and ideas which are standard are there used without elaboration.

Chapter I

(1.1) An Asymptotically Stable Limit Point Theorem

It has been very common to attempt pertubation solutions about a linear solution. For the most part, when this is done in the literature, the convergence of the procedure is simply assumed. In this section, we present a pertubation procedure for a class of problems whose convergence is given in Appendix A.

The basic system with which we shall deal here is given by the equation

$$\frac{dx}{dt} = Ax + G(t)x + H[t, x(t)] + g(t), \qquad x(0) = x_0 \tag{1.11}$$

In this equation, all quantities are continuous functions of time. $x(t)$ and $g(t)$ are n-vectors. A is a constant nxn matrix and $G(t)$ is a time dependent nxn matrix. $H[t, x(t)]$ is a vector whose components are each a polynomial in the components of x. These polynomials are at least of second degree; there is no linear dependence on the components of x. The coefficients in the polynomials are bounded functions of time.

There is a condition on A. It is that the eigenvalues of A all have negative (not zero) real parts. This condition in physical

terms is simply that the system have linear friction. We will come back to a discussion of the terms in the equation later. Now we wish to state a theorem which provides a means of solution of (1.11) under certain circumstances.

We use the notation $\|\Omega\|$ to denote a scalar which equals the largest magnitude of any of the components of Ω whether Ω be a vector or matrix.

Theorem I

For equation (1.11), there exist positive numbers R, G, and g, each > 0 such that if

$$\|x_0\| < R$$

$$\|G(t)\| < G \qquad \text{for all } t$$

$$\|g(t)\| < g \qquad \text{for all } t$$

then the equation can be solved by the following successive approximation scheme:

$x_1(t)$ is the solution to

$$\frac{dx_1}{dt} = Ax_1 + g(t) \quad , \qquad x_1(0) = x_0$$

while for $n \geq 1$

$$\frac{dx_{n+1}}{dt} = Ax_{n+1} + G(t)x_n + H[t, x_n] + g(t),$$

$$x_{n+1}(0) = x_0$$

The successive approximations converge to the solution $x(t)$ uniformly over all time.

This theorem is proved in Appendix A. A more general statement appears at the end of Appendix A with the above theorem as a special case. In practice, the applicability of the hypothesis of the general statement will be too hard to determine. The case presented here will be useful in many applications.

We return to the requirement of the negativeness of the eigenvalues of A Under this condition, the linear equation

$$\frac{dx}{dt} = Ax + g(t) \tag{1.12}$$

whose solution is

$$x(t) = e^{At}x_0 + \int_0^t e^{A(t-\tau)}g(\tau)\,d\tau \tag{1.13}$$

basically controls the situation for the non-linear equation (1.11) subject to the smallness requirements of the theorem.

Let us examine the solution to the linear problem (1.12). For simplicity we suppose the eigenvalues of A not only negative but

also distinct. Then e^{At} can be written in the form

$$e^{At} = B e^{Dt} B^{-1}$$

where D is a diagonal matrix $\begin{pmatrix} \lambda_1 & 0 & \cdot & \cdot & 0 \\ 0 & \lambda_2 & \cdot & \cdot & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & \cdot & \cdot & \lambda_n \end{pmatrix}$, and is time independent.

(1.13) becomes

$$x(t) = B e^{Dt} B^{-1} x_0 + \int_0^t B e^{D(t-\tau)} B^{-1} g(\tau)\, d\tau$$

Define $$y(t) = B^{-1} x(t)$$

$$y_0 = B^{-1} x_0 = B^{-1} x(0) = y(0)$$

$$B^{-1}[g(t)] = h(t)$$

This produces

$$y(t) = e^{Dt} y_0 + \int_0^t e^{D(t-\tau)} h(\tau)\, d\tau \tag{1.14}$$

Or, on a component basis

$$y_i(t) = e^{\lambda_i t} y(0) + \int_0^t e^{\lambda_i (t-\tau)} h_i(\tau)\, d\tau \tag{1.141}$$

From this and the condition $\lambda_i < 0$, we obtain the bound

$$|y_i(t)| < |y_i(0)| + \frac{1}{|\lambda_i|} \sup_{\text{over } t} [|h_i(t)|]$$

By choosing bounds for the initial conditions and the forcing function small enough, we can keep $y(t)$ uniformly bounded as small as we wish over all time. This implies the same thing for $x(t) = B\,y(t)$. Changing (1.12) to

$$\text{(1.15)} \quad \frac{dx}{dt} = Ax + G(t)x + g(t)$$

does not alter the situation as long as $G(t)$ is bounded small enough over all time. This can be proven separately from, but in the same way as, the theorem stated above and is proven in Appendix A.

Now we have the feeling that as long as a linear solution stays small, it ought to be a good first approximation to a non-linear equation. And indeed, the theorem shows this to be true.

For the equation (1.11), the vector point $x(t) = 0$ is a solution for $g(t) = 0$. $x = 0$ is called a limit point. But actually in our problem, it can be shown (Coddington and Levinson, 1955) that under the conditions of the theorem (smallness of x_0 , $G(t)$, and $g(t)$), should $g(t) = 0$ for all $t > t_1$, then each solution tends asymptotically to 0 . The requirement

of $\lambda_i < 0$ is essential here. $x = 0$ is further designated as an asymptotically stable limit point.

The theorem which we have discussed here, proceeds along much the same lines as many stability or asymptotic behavior theorems, (Coddington and Levinson, 1955; Lefschetz, 1957; Birdgland, 1961). It is presented here for a number of reasons. Firstly, the theorem as presented here is in a form particularly useful for application to physical problems. But secondly, asymptotic behavior theorems require the forcing function to go asymptotically to a constant. We do not require this for the theorem here, and hence are led to a somewhat different theorem. This has been done because we have in mind statistical problems. To illustrate this, we go back to the successive approximation scheme.

$$\frac{dx_{n+1}}{dt} = Ax_{n+1} + G(t)x_n + H[t, x_n] + g(t), \qquad x_{n+1}(0) = x_0, \quad n \geq 1 \tag{1.16n}$$

This is equivalent to

$$x_{n+1}(t) = e^{At}x_0 + \int_0^t e^{A(t-\tau)}G(\tau)x_n(\tau)\,d\tau + \int_0^t e^{A(t-\tau)}H[\tau, x_n(\tau)]\,d\tau + \int_0^t e^{A(t-\tau)}g(\tau)\,d\tau \tag{1.17n}$$

The first approximation is

$$x_1(t) = e^{At}x_0 + \int_0^t e^{A(t-\tau)}g(\tau)\,d\tau \tag{1.171}$$

Since $H[\tau, x_n]$ is a polynomial in the components of x_n the successive approximations will produce components of x_{n+1} in terms of products of the components of x_0 and $g(t)$. For instance,

$x_1(0)\, x_3(0)\, g_1(\tau_1)\, g_1(\tau_2)\, g_4(\tau_3)\, g_5(\tau_4)$ could be one such term.

Products like $x_i(t)\, x_j(t+\lambda)$ are approximated uniformly over time and over all properly bounded initial conditions and forcing functions. Hence the n^{th} approximation, taken over an ensemble average, gives $\overline{x_i(t)_n\, x_j(t)_n}$ in terms of ensemble averages like $\overline{x_1(0)\, x_3(0)\, g_1(\tau_1)\, g_1(\tau_2)\, g_4(\tau_3)\, g_5(\tau_4)}$.

$\overline{x_i(t)_n\, x_j(t+\lambda)_n}$ converges to $\overline{x_i(t)\, x_j(t+\lambda)}$ uniformly over all time and all sufficiently bounded ensembles of initial conditions and forcing function.

In this way, one successively approximates the output statistics in terms of the input statistics. A special but important situation presents itself immediately. Suppose that $G(t) = 0$ and $H[t; x]$ is not explicitly a function of t. $H[t, x] = H(x)$.

Suppose further that all input statistics are stationary. Then all output statistics will also be stationary as $t \longrightarrow \infty$. That is, a statistic like $\overline{x_i(t)\, x_j(t+\lambda)}$ will go to a definite limit, dependent on λ, as $t \longrightarrow \infty$.

(1.2) Viscous Incompressible Two-Dimensional Water Waves

As an example of the possible application of the foregoing theorem, we here set up an approximate viscous incompressible water wave theory.

There is however a serious approximation into which we are forced right from the beginning. It is that we will try to describe the problem in terms of a scalar potential. This procedure would of course be correct in the inviscid case. In fact, special stress combinations will produce curl free motion, and in these cases the problem can be formulated by the scalar potential. It is felt by the author that there are difficulties with the classical viscous linear problem (Lamb, 1932; Handbuch der Physik, 1960). The classical methods do not give the same results as those obtained here for the linear problem forced by a special curl free motion producing stress configuration. Hence at a later time, it is hoped that the classical method for the linear problem, as well as the procedure below, will be checked by machine computation.

The defense of an approximation such as we use here must ultimately rest on the fact that it can actually be carried out, and that at least the form of the results is more or less what one would expect. It will be easily seen that as the viscosity goes to zero and the amplitudes become infinitessimal, the procedure yields the correct linear inviscid model.

Let the depth of the water be infinite at $z = -\infty$. $z = 0$ is the mean value of the surface. The lateral dimension is x. There is no dependence on y.

If u stands for the velocity along the x-axis and w the velocity along the z-axis, the incompressibility of the fluid introduced into the equation of continuity gives

$$\frac{\partial u}{\partial x} + \frac{\partial w}{\partial z} = 0 \quad . \tag{1.21}$$

The force equations are

$$\frac{\partial u}{\partial t} + u\frac{\partial u}{\partial x} + w\frac{\partial u}{\partial z} = F_x - \frac{1}{\rho}\frac{\partial P}{\partial x} + \nu\nabla^2 u \tag{1.22a}$$

and

$$\frac{\partial w}{\partial t} + u\frac{\partial w}{\partial x} + w\frac{\partial w}{\partial z} = F_z - \frac{1}{\rho}\frac{\partial P}{\partial z} + \nu\nabla^2 w \tag{1.22b}$$

where P is pressure, ρ is density, and F_x and F_z are body forces in the x and z direction respectively. ν is the coefficient

of viscosity and assumed small. In particular, these equations hold not only throughout the fluid, but also at the surface.

The kinematic condition at the wave surface requires that a particle on the surface remains on the surface. If $\eta(x, t)$ is the deviation from the mean of the amplitude of the surface motion

$$\frac{\partial \eta}{\partial t} + u \frac{\partial \eta}{\partial x} - w = 0 . \tag{1.23}$$

There are two other conditions on the problem. They are the stress conditions normal and tangential to the boundary surface. These will be determined ultimately from the stress-rate of strain relations.

$$p_{zz} = -p + 2\rho\nu \frac{\partial w}{\partial z} \tag{1.24a}$$

$$p_{xz} = \rho\nu \left(\frac{\partial w}{\partial x} + \frac{\partial u}{\partial z} \right) \tag{1.24b}$$

$$p_{xx} = -p + 2\rho\nu \frac{\partial u}{\partial x} \tag{1.24c}$$

To obtain the required relations, we pick an arbitrary point on the wave surface and draw two sets of rectangular axes each of whose origins is located at the point. The first set of axes are horizontal and vertical x-z-axes respectively. The second set of axes are drawn such that the x'-axis is tangent to, and the z'-axis is normal to the wave surface at the chosen point. If θ is the angle between

the x- and x'-axes, then

$$\frac{\partial \eta}{\partial x} = \tan\Theta \approx \Theta$$

for small enough angles. η itself is not allowed large. The condition $\tan\Theta \approx \Theta$ is a smoothness condition on η.

Remembering that $P_{xz} = P_{zx}$, and using the transformation relations of Cartesian Tensor Analysis (Jeffreys, 1931), we have after neglecting terms with factors of squared and higher orders in Θ,

(1.25a) $$P_{z'z'} = P_{zz} - 2\frac{\partial \eta}{\partial x} P_{xz}$$

(1.25b) $$P_{z'x'} = P_{zx} + \frac{\partial \eta}{\partial x}\left(P_{zz} - P_{xx}\right)$$

We still do not quite have the relations we want. By expanding in a Taylor series and retaining only through second order non-linearity, (1.25a,b) become respectively

(1.26a) $$P_{z'z'} = -P\big|_{z=\eta} + 2\rho\nu \frac{\partial w}{\partial z}\Big|_{z=0} + 2\rho\nu \frac{\partial^2 w}{\partial z^2}\Big|_{z=0}^{\eta} - 2\rho\nu \frac{\partial \eta}{\partial x}\left[\frac{\partial w}{\partial x} + \frac{\partial u}{\partial z}\right]_{z=0}$$

$$P_{z'x'} = \rho\nu\left(1 + \eta\frac{\partial}{\partial z}\right)\left(\frac{\partial w}{\partial x} + \frac{\partial u}{\partial z}\right)\Big|_{z=0}$$

(1.26b)

$$+ 2\rho\nu\frac{\partial \eta}{\partial x}\left(\frac{\partial w}{\partial z} - \frac{\partial u}{\partial x}\right)\Big|_{z=0}$$

(1.26a,b) are the required stress relations.

We will first take up the case of zero boundary stress. From (1.26a) we obtain the pressure at the surface in terms of the velocity components and the amplitude η .

It is assumed that neither η nor the velocity components or their derivatives are large. Also ν is very very small. Our major approximation then is that $P_{z'x'}$ will be not quite but nearly equal to zero. We will come back to this later; but in the immediately following discussion, (1.26b) will be dispensed with entirely.

At this point we introduce the scalar velocity potential ϕ

$$u = -\frac{\partial \phi}{\partial x} \quad , \qquad w = -\frac{\partial \phi}{\partial z}$$

The force equations (1.22a,b) become

$$\text{grad}\left[\frac{\partial \phi}{\partial t} - \frac{1}{2}\left(\frac{\partial \phi}{\partial x}\right)^2 - \frac{1}{2}\left(\frac{\partial \phi}{\partial z}\right)^2\right] =$$

$$= \text{grad}\left[g\eta + \frac{1}{\rho}P\right].$$

To obtain this we have used $\vec{F} = -\text{grad}\,[g\eta]$ while $\nabla^2\phi = 0$ from the incompressibility of the fluid.

This relation holds throughout the fluid, and in particular at the surface. With the aid of (1.26a) we obtain through second order non-linearity

$$\left.\frac{\partial\phi}{\partial t}\right|_{z=0} + \left.\frac{\partial^2\phi}{\partial t\,\partial z}\right|_{z=0}\eta - \frac{1}{2}\left[\left(\frac{\partial\phi}{\partial x}\right)^2 + \left(\frac{\partial\phi}{\partial z}\right)^2\right]_{z=0}$$

$$- g\eta + 2\nu\left.\frac{\partial^2\phi}{\partial z^2}\right|_{z=0} + 4\nu\left.\frac{\partial^3\phi}{\partial z^3}\right|_{z=0}\eta$$

$$- 4\nu\left.\frac{\partial^2\phi}{\partial x\,\partial z}\right|_{z=0}\frac{\partial\eta}{\partial x} = 0 \qquad (1.28)$$

In (1.28), consider the operator defined by

$$\left.\frac{\partial\phi}{\partial t}\right|_{z=0} + \left.\frac{\partial^2\phi}{\partial t\,\partial z}\right|_{z=0}\eta = \left(I + \eta\,\frac{\partial}{\partial z}\right)\left.\frac{\partial\phi}{\partial t}\right|_{z=0}$$

Here I is simply the identity operator. Now if the operator $\left[\eta\,\frac{\partial}{\partial z}\right]$ is in some sense small enough, then the inverse of the operator $\left(I + \eta\,\frac{\partial}{\partial z}\right)$, can be obtained as (Friedman, 1956)

$$\left(I + \eta \frac{\partial}{\partial z}\right)^{-1} = I - \eta \frac{\partial}{\partial z} + \eta^2 \frac{\partial^2}{\partial z^2} - \cdots$$

Without hesitation, we make the smoothness assumption that $\left[\frac{\partial}{\partial z}\right]$ is a bounded operator. Then since η has already been supposed very small, the operator $\left(I + \eta \frac{\partial}{\partial z}\right)$ can be inverted as described.

Again through second order, (1.28) becomes

$$\text{(1.29)} \quad \frac{\partial \phi}{\partial t}\Big|_{z=0} - g\eta + 2\nu \frac{\partial^2 \phi}{\partial z^2}\Big|_{z=0} - \frac{1}{2}\left[\left(\frac{\partial \phi}{\partial x}\right)^2 + \left(\frac{\partial \phi}{\partial z}\right)^2\right]_{z=0} + 2\nu \frac{\partial^3 \phi}{\partial z^3}\Big|_{z=0}\eta - 4\nu \frac{\partial^2 \phi}{\partial x \partial z}\Big|_{z=0}\frac{\partial \eta}{\partial x} = 0$$

In terms of the potential ϕ, the kinematic condition (1.23) through second order is given by

$$\text{(1.30)} \quad \frac{\partial \eta}{\partial t} + \frac{\partial \phi}{\partial z}\Big|_{z=c} + \frac{\partial^2 \phi}{\partial z^2}\Big|_{z=0}\eta - \frac{\partial \phi}{\partial x}\Big|_{z=0}\frac{\partial \eta}{\partial x} = 0$$

We have now only to set

$$\phi(t,x,z) = \int_{-\infty}^{\infty} \phi_k(t)\, e^{ikx}\, e^{|k|z}\, dk$$

$$\eta(t,x) = \int_{-\infty}^{\infty} \eta_k(t)\, e^{ikx}\, dk$$

and substitute in (1.29) and (1.30). Notice that $\nabla^2\phi(t,x,z) = 0$ which is the incompressibility requirement.

(1.29) goes over to

$$\text{(1.31)}\quad \frac{d\phi_k}{dt} - g\eta_k + 2\nu k^2\phi_k + \frac{1}{2}\int_{-\infty}^{\infty} [\ell(k-\ell) - |\ell(k-\ell)|]\,\phi_\ell\,\phi_{k-\ell}\,d\ell + 2\nu\int_{-\infty}^{\infty} |\ell^3|\,\phi_\ell\,\eta_{k-\ell}\,d\ell - 4\nu\int_{-\infty}^{\infty} [|\ell^3| - \ell k|\ell|]\,\phi_\ell\,\eta_{k-\ell}\,d\ell = 0$$

And (1.30) transforms into

$$\text{(1.32)}\quad \frac{d\eta_k}{dt} + |k|\,\phi_k + \int_{-\infty}^{\infty} \ell^2\phi_\ell\,\eta_{k-\ell}\,d\ell + \int_{-\infty}^{\infty} \ell(k-\ell)\,\phi_\ell\,\eta_{k-\ell}\,d\ell = 0$$

(1.30) and (1.32) are in the required form for the theorem to apply, except that they represent an infinite number of equations. For the eigenvalues of the linear part are

$$\lambda = -\nu k^2 \pm \frac{\sqrt{(2\nu k^2)^2 - 4|gk|}}{2}$$

. The cut-off frequency for k is not large. As discussed in section (0.5) of the Introduction, for any given numerical problem, only a finite number of equations would be used. Hence we can suppose that for any practical purpose, the successive approximation scheme given by Theorem I will converge uniformly over all time and sufficiently small $\eta(0, x)$ and $\phi(0, x, z)$.

It is to be noted that in any practical problem, we must be able to find a $k_1 > 0$ such that setting $\phi_k = 0$ for $|k| > k_1$ is a reasonable approximation. Then since

$$\left|\frac{\partial \phi}{\partial z}\right| = \left|\int_{-\infty}^{\infty} |k| \phi_k e^{ikx} e^{|k|z} dk\right|$$

$$\leq \int_{-k_1}^{k_1} |k| |\phi_k| dk$$

we will have $\frac{\partial \phi}{\partial z}$ uniformly bounded. This substantiates the inversion of the operator $\left(I + \eta \frac{\partial}{\partial z}\right)$ as carried out above, and allows λ to be approximately $-\nu k^2 \pm i\sqrt{|gk|}$.

Still neglecting for a moment the boundary condition (1.26b), let us take $P_{3'3'}$ small but not identically zero in (1.26a), and investigate the linear problem.

The linearized equivalent of (1.31) and (1.32) with

$P_{3'3'} \approx P_{33}(x)$ included are

$$\frac{d\phi_k}{dt} + 2\nu k^2 \phi_k - g\eta_k = -\frac{1}{\rho}\int_{-\infty}^{\infty} P_{33}(x) e^{-ikx} dx \tag{1.33}$$

$$\frac{d\eta_k}{dt} + |k|\phi_k = 0 \tag{1.34}$$

Suppose that we solved this linear problem, and then substituted the results into the linear approximation for the boundary condition (1.26b). It is the condition which we have neglected so for, and would lead to the requirement on $P_{3'x'} \approx P_{3x}$ that

$$P_{3x} = \rho\nu\left(\frac{\partial w}{\partial x} + \frac{\partial u}{\partial 3}\right)$$

Hence for any small enough P_{33}, there exists a P_{3x} such that the linearized motion is curl free and given by a scalar potential. But the same can be said for the non-linear problem. We equate the left hand side of (1.31) to $-\frac{1}{\rho} P_{3'3'}(k)$ and

solve this equation together with (1.32) by means of Theorem I. Substitution of the result into (1.26b) gives the condition on $P_{3'x'}$ such that one has potential motion, and hence has actually solved the problem.

Although only a very restricted class of problems can be handled in this way, the method may furnish some clues about the non-linear relation between the wave spectrum and the input energy spectrum.

Chapter II

(2.1) An Initial Value Problem

We will now turn our attention away from pertubation procedures and obtain a truly non-linear result. The problem under consideration is

$$\frac{dx}{dt} = f(t,x) \quad , \qquad x(0) = x_c \tag{2.11}$$

where $f(t,x)$ is an n-vector. The components of f are each analytic functions of $t, x_1, \cdots, x_n$ over some rectangle R,

$$|t| < T$$

$$|x_i| < R_i \quad , \quad i = 1, 2, \cdots, n$$

This means that for any fixed point in R each component of f is an analytic function in each variable taken one at a time.

Note that we require in addition $f(t,0) = 0$. There is no forcing function for (2.11).

For the solution to (2.11), we will obtain a representation theory. The proof depends upon two major known theorems. The proof itself is an immediate result of the statement of these theorems.

The first theorem needed is due to Poincaré (Lefschetz, 1957). One result of this theorem as applied to (2.11) is as follows:

Let f and R be defined as above. Let B be a rectangle of initial conditions defined by

$$0 \leq |x(0)_i| < B_i \leq R_i \quad , \qquad i = 1, 2, \cdots n$$

We suppose that the B_i have been so chosen that if $x(0)$ originates in B, then the solution for $0 \leq t < d \leq T$ remains in R. Then about each point in B, there exists a neighborhood $N(x_0)$ such that each component of the solution is an analytic function of

$t, x_{0_1}, x_{0_2}, \cdots, x_{0_n}$ for $0 \leq t < d$.

The second theorem we require is a result from the theory of several complex variables (Bochner and Martin, 1948). First let us define the rectangle $d \times B$ as

$$0 \leq |t| < d$$
$$0 \leq |x_i| < B_i \quad , \qquad i = 1, 2, \cdots, n$$

Then we can state the theorem as:

Theorem II

If a function of several complex variables is analytic in each variable separately at each point of a rectangle such as $d \times B$ above, then the function is analytic simultaneously in all variables over the whole rectangle.

The theorem means that the function can be represented as a multiple power series in the $n+1$ variables. Each component of the solution to (2.11) is given by

$$(2.12) \qquad x_i(t) = \sum_{pq} C^i_{p_1 \cdots p_n q} x_1(0)^{p_1} \cdots x_n(0)^{p_n} t^q$$

This series converges absolutely and uniformly over any closed rectangle such as

$$0 \le |t| \le c < d$$

$$0 \le |x_i| \le A_i < B_i \ , \quad i = 1, 2, \cdots, n$$

contained in $d \times B$.

An analytic function of several complex variables has many of the properties of a function of a single complex variable. For instance, the partial derivative with respect to any variable can be found by changing the order of differentiation and summation. The resulting series is also absolutely and uniformly convergent over the same closed rectangle as the original series. The same holds for any number of partial derivatives. Further one can sum over all indices simultaneously, or sum over one index completely first.

Let us return to (2.12) as a form of solution for (2.11). Since order of operations like summation and differentiation is immaterial, and $x(0) = 0$ implies $x(t) = 0$, we can represent the solution to (2.11) as

$$(2.13)\qquad x_i(t) = \sum_{j=1}^{n} C_j^i(t)\, x_j(0) + \sum_{j_1=1}^{n}\sum_{j_2=1}^{n} C_{j_1 j_2}^i(t)\, x_{j_1}(0)\, x_{j_2}(0)$$

$$+ \sum_{j_1=1}^{n}\sum_{j_2=1}^{n}\sum_{j_3=1}^{n} C_{j_1 j_2 j_3}^i(t)\, x_{j_1}(0)\, x_{j_2}(0)\, x_{j_3}(0)$$

$$+ \cdots$$

At $t = 0$ this becomes

$$x_i(0) = \sum_{j=1}^{n} C_j^i(0)\, x_j(0) + \sum_{j_1=1}^{n}\sum_{j_2=1}^{n} C_{j_1 j_2}^i(0)\, x_{j_1}(0)\, x_{j_2}(0)$$

$$+ \sum_{j_1=1}^{n}\sum_{j_2=1}^{n}\sum_{j_3=1}^{n} C_{j_1 j_2 j_3}^i(0)\, x_{j_1}(0)\, x_{j_2}(0)\, x_{j_3}(0) + \cdots$$

Since the relation must hold for all x in B. We obtain

$$(2.14)\qquad C_i^i(0) = 1 \quad ; \quad C_j^i(0) = 0 \quad , \quad i \neq j$$

$$C_{j_1 j_2}^i(0) = C_{j_1 j_2 j_3}^i(0) = \cdots = 0$$

When we solve for the C's, (2.14) will be the initial conditions.

Now in (2.11) since $f(t, x)$ is analytic, and $f(t, 0) = 0$ we can also represent $f_i(t, x)$ as

$$(2.15)\quad f_i(t,x) = \sum_{j=1}^{n} f_j^i(t)\, x_j(t) + \sum_{j_1=1}^{n}\sum_{j_2=1}^{n} f_{j_1 j_2}^i(t)\, x_{j_1}(t)\, x_{j_2}(t) + \sum_{j_1=1}^{n}\sum_{j_2=1}^{n}\sum_{j_3=1}^{n} f_{j_1 j_2 j_3}^i(t)\, x_{j_1}(t)\, x_{j_2}(t)\, x_{j_3}(t) + \cdots$$

We use (2.15) for $f(t,x)$ and (2.13) with the conditions (2.14) for $x(t)$ in (2.11). The resulting equation must hold for all $x(0)$ in B. This determines that

$$(2.151)\quad \frac{d}{dt} C_k^i(t) = \sum_{j=1}^{n} f_j^i(t)\, C_k^j(t), \quad k \text{ fixed}, \quad i = 1, 2, \cdots, n$$

$$C_k^i(0) = 0, \quad i \neq k$$

$$C_k^k(0) = 1$$

$$(2.152)\quad \frac{d}{dt} C_{k_1,k_2}^i = \sum_{j=1}^{n} f_j^i(t)\, C_{k_1,k_2}^j + \sum_{j_1=1}^{n}\sum_{j_2=1}^{n} f_{j_1 j_2}^i(t)\, C_{k_1}^{j_1}\, C_{k_2}^{j_2}, \quad k_1, k_2 \text{ fixed}, \quad i = 1, 2, \cdots, n$$

$$C_{k_1,k_2}^i(0) = 0$$

$$\frac{d}{dt} C^{i}_{k_1 k_2 k_3} = \sum_{j=1}^{n} f^{i}_{j}(t)\, C^{j}_{k_1 k_2 k_3}$$

(2.153)

$$+ 2 \sum_{j_1=1}^{n} \sum_{j_2=1}^{n} f^{i}_{j_1 j_2}(t)\, C^{j_1}_{k_1 k_2}\, C^{j_2}_{k_3}$$

$$+ \sum_{j_1=1}^{n} \sum_{j_2=1}^{n} \sum_{j_3=1}^{n} f^{i}_{j_1 j_2 j_3}(t)\, C^{j_1}_{k_1}\, C^{j_2}_{k_2}\, C^{j_3}_{k_3}$$

where k_1 and k_2 and k_3 are fixed; $\quad i = 1, 2, \cdots, n$;

$C^{i}_{k_1 k_2 k_3}(0) = 0$.

$$\frac{d}{dt} C^{i}_{k_1 \cdots k_n} = \sum_{j=1}^{n} f^{i}_{j}(t)\, C^{j}_{k_1 \cdots k_n} + g^{i}_{k_1 \cdots k_n}(t) ,$$

(2.15n) $n > 1$

$k_1, \cdots, k_n$ fixed,

$i = 1, 2, \cdots, n$

$C^{i}_{k_1 \cdots k_n}(0) = 0$

where $g^{i}_{k_1 \cdots k_n}(t)$ is a forcing function which depends at most on C's already determined.

We wish to emphasize a feature of (2.15n). In terms of the unknown $C^{i}_{k_1 \cdots k_n}$ the equation is linear. We can rewrite

(2.15n) as an operator equation

$$L\left(C^{i}_{k_1,\dots k_n}(t)\right) = g^{i}_{k_1,\dots k_n}(t) \quad , \quad i = 1,\dots, n \tag{2.16n}$$

$$C^{i}_{k_1,\dots k_n}(0) = 0$$

where L is a linear operator. The important thing is that L does not depend on n . Hence in any given problem, one determines the inverse to L only once, and then operates each successive step with L^{-1} on the already known forcing function. When the coefficients of the linear term in the expansion for f given by (2.15) are constants, the inverse operator is particularly simple. In this case, (2.16n) becomes

$$\frac{d}{dt} C^{i}_{k_1,\dots k_n} = \sum_{j=1}^{n} f^{i}_{j} C^{j}_{k_1,\dots k_n} + g^{i}_{k_1,\dots k_n}(t) \quad , \tag{2.17n}$$

$n>1$

$k_1, \dots, k_n$ fixed

$i = 1, 2, \dots, n$

$$C^{i}_{k_1,\dots k_n}(0) = 0$$

The solution to (2.17n) in vector form is

$$C_{k_1,\dots k_n}(t) = \int_{0}^{t} e^{f(t-\tau)}\, g_{k_1,\dots k_n}(\tau)\, d\tau \tag{2.18n}$$

$n>1$

where $C_{k_1 \cdots k_n}$ and $g_{k_1 \cdots k_n}$ are n-vectors whose components are $C^i_{k_1 \cdots k_n}$ and $g^i_{k_1 \cdots k_n}$ respectively; f is a matrix whose components are f^i_j.

(2.2) Conservative Two-Dimensional Vorticity

We here follow Lorenz (1953). The motion is assumed incompressible and confined to the $x - y$ plane. Hence the velocity components can be obtained from a stream function ψ.

Velocity in the x direction $= u = -\frac{\partial \psi}{\partial y}$

Velocity in the y direction $= v = \frac{\partial \psi}{\partial x}$.

Taking the curl of the force equations (1.22a,b) we obtain the vorticity equation

$$\nabla^2 \frac{\partial \psi}{\partial t} = \frac{\partial \psi}{\partial y} \cdot \nabla^2 \frac{\partial \psi}{\partial x} - \frac{\partial \psi}{\partial x} \nabla^2 \frac{\partial \psi}{\partial y} \tag{2.21}$$

It is supposed that there are no boundaries. The equation holds over the whole plane. Hence we represent ψ by

$$\psi(t,x,y) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \psi_{mn}(t) e^{imx} e^{iny} \, dm \, dn$$

This leads to

(2.221)

$$\frac{d\,\psi_{mn}}{dt} = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\left(\frac{mq-np}{m^2+n^2}\right)(p^2+q^2)\,\psi_{pq}\,\psi_{m-p,n-q}\,dp\,dq$$

Since $\psi(t,x,y)$ is a real function, $\overline{\psi_{mn}} = \psi_{-m,-n}$

where the bar denotes a complex conjugate. Taking the complex conjugate of (2.221)

(2.222)

$$\frac{d\,\overline{\psi_{mn}}}{dt} = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\left(\frac{mq-np}{m^2+n^2}\right)(p^2+q^2)\,\psi_{-p,-q}\,\psi_{p-m,q-n}\,dp\,dq$$

The total energy in the field is

$$E = \frac{1}{2}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\left\{[u(x,y)]^2+[v(x,y)]^2\right\}dx\,dy$$

which by Parseval's Theorem is equal to

$$\frac{1}{2}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}(m^2+n^2)\,\psi_{mn}\,\overline{\psi_{mn}}\,dm\,dn$$

The change of E with time is zero since the system is conservative.

$$(2.23)\quad \frac{dE}{dt} = 0$$

$$= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}(m^2+n^2)\left[\psi_{-m,-n}\frac{d\psi_{mn}}{dt} + \psi_{mn}\frac{d}{dt}\psi_{-m,-n}\right]dm\,dn$$

We shall however, be interested in the quantity

$$E_{mn}(t) = (m^2+n^2)\,\psi_{mn}\,\overline{\psi_{mn}} = (m^2+n^2)\,\psi_{mn}\,\psi_{-m,-n}$$

which is the energy density over time.

Let us return to the fundamental equation (2.221). This equation is of the form required by Theorem II, except that it is an infinite system. We again adopt the attitude that in practice, a finite system approximation is adequate. Hence we solve (2.221) by the recursive scheme of Theorem II.

So as not to confuse letters, let $\psi_{mn}(0) = A_{mn}$.

Then the scheme indicates the representation for $\psi_{mn}(t)$

$$(2.24)\quad \psi_{mn}(t) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} C^{mn}_{pq}(t)\,A_{pq}\,dp\,dq$$

$$+\int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty} C^{mn}_{p_1q_1p_2q_2}(t)\,A_{p_1q_1}A_{p_2q_2}\,dp_1\cdots dq_2+\cdots$$

The initial conditions on the C's are obtained from the requirement that $\psi_{mn}(0) = A_{mn}$. Hence

(2.25) $$C^{mn}_{pq}(0) = \delta(m,p)\,\delta(n,q)$$

$$C^{mn}_{p_1q_1p_2q_2}(0) = 0$$

$$C^{mn}_{p_1q_1p_2q_2p_3q_3}(0) = 0 \qquad \text{etc.}$$

Putting this into (2.221) gives first

(2.261) $$\frac{d}{dt}\,C^{mn}_{pq}(t) = 0$$

Hence

(2.271) $$C^{mn}_{pq}(t) = \delta(m,p)\,\delta(n,q)$$

The next equation obtained is

(2.62) $$\frac{d}{dt}\,C^{mn}_{p_1q_1p_2q_2}(t) =$$

$$= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \frac{(ms-nr)(r^2+s^2)}{(m^2+n^2)}\,C^{rs}_{p_1q_1}\,C^{m-r,n-s}_{p_2q_2}\,dr\,ds$$

Taking account of the initial condition $C^{mn}_{p_1q_1p_2q_2}(0) = 0$

and (2.271), we obtain as solution to (2.262)

(2.272) $$C^{mn}_{p_1q_1,p_2q_2}(t) =$$

$$= \frac{(mq_1 - np_1)(p_1^2+q_1^2)}{(m^2+n^2)} \delta(m, p_1+p_2)\,\delta(n, q_1+q_2)\,t$$

We proceed in the same way to

(2.263) $$\frac{d}{dt} C^{mn}_{p_1q_1,p_2q_2,p_3q_3}(t) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} dr\,ds \cdot$$

$$\cdot \left\{ \frac{(ms-nr)(r^2+s^2)}{(m^2+n^2)} \left[C^{rs}_{p_1q_1} C^{m-r,n-s}_{p_2q_2,p_3q_3} + C^{rs}_{p_1q_1,p_2q_2} C^{m-r,n-s}_{p_3q_3} \right] \right\}$$

with the initial condition that $C^{mn}_{p_1\cdots q_3}(0) = 0$.

The solution to this is (2.273).

(2.273) $$C^{mn}_{p_1\cdots q_3}(t) =$$

$$= \frac{t^2}{2} \left\{ \frac{(mq_1 - np_1)(p_1^2+q_1^2)(p_2^2+q_2^2)\left[(m-p_1)q_2 - (n-q_1)p_2\right]}{(m^2+n^2)\left[(m-p_1)^2 + (n-q_1)^2\right]} \cdot \right.$$

$$\left. \cdot\, \delta(m-p_1, p_2+p_3)\,\delta(n-q_1, q_2+q_3) \right\}$$

It is easily seen that all of the C's will combine to give, for fixed initial conditions, a Taylor series type expansion over time. This is because the original equation (2.221) had no linear term on the right hand side.

Now the scheme (2.24) will converge uniformly over t and A_{pq} for bounded t and bounded A_{pq} . (This is strictly true when only a finite number of A's are used). But a Taylor series does not often converge rapidly. Nevertheless, the scheme (2.24), when used to obtain the energy density, will perhaps be useful in studying the stability of statistically described flows over a small but finite time interval. The reader is again refered to Lorenz (1953) who initiated there the instantaneous stability problem.

The convergence problem being as it is above, we turn in the next section to a dissipative system where this difficulty is overcome in a definite sense.

(2.3) Non-Conservative Turbulence as an All-Time Problem

We envision a two-dimensional viscous fluid extending over

$$\begin{bmatrix} -\infty \leq x \leq \infty \\ -\infty \leq y \leq \infty \end{bmatrix}$$

(The three dimensional problem would follow in the same way as below).

The basic force equations have been given in section (1.2) as equations (1.22a,b). We will use

$$\text{(2.31a)}\quad \frac{\partial u}{\partial t} + u\frac{\partial u}{\partial x} + w\frac{\partial u}{\partial z} = -\frac{1}{\rho}\frac{\partial p}{\partial x} + \nu\nabla^2 u$$

$$\text{(2.31b)}\quad \frac{\partial w}{\partial t} + u\frac{\partial w}{\partial x} + w\frac{\partial w}{\partial z} = -\frac{1}{\rho}\frac{\partial p}{\partial z} + \nu\nabla^2 w$$

The quantities in (2.31a,b) are defined as in (1.22a,b). The equations (2.31a,b) are the same as (1.22a,b), except that the external field has been deleted.

In addition, we still have the equation of continuity.

$$\text{(2.32)}\quad \frac{\partial u}{\partial x} + \frac{\partial w}{\partial z} = 0$$

This is the same as equation (1.21).

By manipulating (2.31a,b), and using (2.32), we obtain

$$\text{(2.33)}\quad -\frac{1}{\rho}\nabla^2 p = \left(\frac{\partial u}{\partial x}\right)^2 + \left(\frac{\partial w}{\partial z}\right)^2 + 2\frac{\partial u}{\partial z}\frac{\partial w}{\partial x}$$

At this point we introduce Fourier coefficients.

$$p = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} p_{mn}\, e^{imx}\, e^{inz}\, dm\, dn$$

$$u = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} U_{mn}\, e^{imx}\, e^{i\cdot nz}\, dm\, dn$$

A similar expression obtains for W .

There are no boundaries in the problem. Hence the operator is easily inverted in the Fourier domain. (2.33) implies

$$(2.34)\quad -\frac{1}{\rho} P_{mn} = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \frac{r(m-r)}{(m^2+n^2)} U_{rs} U_{m-r,n-s}\, dr\,ds + \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \frac{s(n-s)}{(m^2+n^2)} W_{rs} W_{m-r,n-s}\, dr\,ds + 2\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \frac{s(m-r)}{(m^2+n^2)} U_{rs} W_{m-r,n-s}\, dr\,ds$$

Equations (2.31a,b) go over to

$$(2.35a)\quad \frac{dU_{mn}}{dt} + \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \left[ir U_{rs} U_{m-r,n-s} + is U_{rs} W_{m-r,n-s}\right] dr\,ds = \frac{-im}{\rho} P_{mn} - \nu(m^2+n^2) U_{mn}$$

$$(2.35b)\quad \frac{dW_{mn}}{dt} + \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \left[ir U_{m-r,n-s} + is W_{m-r,n-s}\right] W_{rs}\, dr\,ds = -\frac{in}{\rho} P_{mn} - \nu(m^2+n^2) W_{mn}$$

P_{mn} is expressed in terms of U_{mn} and W_{mn} through (2.34). Taking (2.34) into account, the linear approximation

to (2.35a,b) is

$$\frac{dU_{mn}}{dt} = -\nu(m^2+n^2)U_{mn} \tag{2.36a}$$

$$\frac{dW_{mn}}{dt} = -\nu(m^2+n^2)W_{mn} \tag{2.36b}$$

The solution to (2.36a,b) is of course

$$U_{mn}(t) = e^{-\nu(m^2+n^2)t}U_{mn}(0)$$

$$W_{mn}(t) = e^{-\nu(m^2+n^2)t}W_{mn}(0)$$

We will not do so, but could here go through the same procedure as in section (2.2), and indicated by Theorem II.

It is perhaps possible to see by inspection however, that unlike the preceding problem, we will here not get (for fixed initial conditions) power series in time. Rather, we will obtain a series whose terms are constants or damping exponentials. Hence (2.35a,b) with (2.34) are diffusion type equations. This is because we have deleted any external field such as the gravity field.

It might at first be supposed that because we have to deal only with damping exponentials, that the representation scheme of Theorem II would in some way converge uniformly over all time. This is unfortunately not true in general. There exist not at all pathological systems such that

with initial conditions fixed, each term in the representation series will be a damping exponential; but the series will not converge as $t \to \infty$.

However, in our present case, we have a special situation. The non-linear terms are conservative, while the linear term is dissipative. The total energy continuous decreases and asymptotically goes to zero. This implies that for arbitrary $\epsilon > 0$, there exists a $\delta(\epsilon) > 0$ such that

$$\max_{m,n} \left[|U_{mn}(t_0)| \right] < \delta(\epsilon)$$

and $$\max_{m,n} \left[|W_{mn}(t_0)| \right] < \delta(\epsilon)$$

implies $|U_{mn}(t)| < \epsilon$ for all $t > t_c$,

and $|W_{mn}(t)| < \epsilon$ for all $t > t_0$.

In addition, $|U_{mn}(t)|$ and $|W_{mn}(t)| \to 0$.

It is shown in Appendix B, that there then exists a $T(R, \epsilon)$ such that for any $R > 0$ and $\epsilon > 0$

$$\max_{m,n} \left[|U_{mn}(0)|, |W_{mn}(0)| \right] \leq R$$

implies $$\max_{m,n} \left[|U_{mn}(t)|, |W_{mn}(t)| \right] < \epsilon$$

for all $t > \tau(R, \epsilon)$.

Since for any R and $\epsilon/2$ and $\tau(R, \epsilon/2)$, there exists an $\epsilon/2$ approximation to the solution uniformly over all initial conditions bounded by R and t bounded by τ , setting the approximate solution equal to zero for $t > \tau$ gives us a representation scheme which converges uniformly over all time at a rate dependent only on the bound on the initial conditions.

Once again, the special property of the above system is that any bounded initial starting configuration damps to zero asymptotically. Other systems satisfying the conditions of Theorem II, while at the same time displaying this property, can be handled in the same way.

Batchelor (1960) has suggested that in homogeneous turbulence, the velocity correlations become Gaussian as $t \longrightarrow \infty$ regardless of the form of the initial distributions. We mention this as suggestive of at least one experiment to be undertaken with the above results.

Chapter III

(3.1) Equations with Forcing Functions Over an Arbitrary Large But Finite Time Interval

In the preceding chapter, the equations had analytic right-hand sides, but were without forcing functions. This chapter retains the condition of analyticity while allowing the equations to possess a forcing term. Our basic system in this chapter will take the form

$$\frac{dx}{dt} = Ax + f(x,t) + g(t) \tag{3.11}$$

where A is an nxn matrix, and x, f and g are n-vectors. f has no linear dependence on x. Also, if $f(0,t)$ were not equal to zero, we could incorporate this term of f into g. Hence we explicitly assume $f(0,t) = 0$. On a component basis, we can write f as

$$f_i(t,x) = f_{j_1 j_2}(t)\, x_{j_1} x_{j_2} + f_{j_1 j_2 j_3}(t)\, x_{j_1} x_{j_2} x_{j_3} + \cdots + f_{j_1 \cdots j_n}(t)\, x_{j_1} \cdots x_{j_n} \tag{3.12i}$$

wherein we have employed the summation convention that any repeated index is to be summed unless otherwise specified. This convention will be in effect throughout the chapter. We have supposed f to have a finite number of terms for convenience.

We are going to introduce a representation scheme which is to hold over any finite time and for at least some large class of forcing functions. The scheme is essentially that of Volterra Functionals. We have already made mention of Volterra Functionals in the Introduction.

We will proceed formally at first. Statements about convergence will be made later. The proofs are mathematical and left to Appendix C.

Consider the simple first degree scalar equation

$$\frac{dx}{dt} + ax = bx^2 + g(t) \tag{3.13}$$

(3.13) is equivalent to

$$x(t) = e^{-at}x_0 + \int_0^t e^{-a(t-\tau)}[bx^2(\tau) + g(\tau)]d\tau \tag{3.14}$$

If $x(t)$ is representable by Volterra Functionals

$$x(t) = K_0(t) + \int_0^t K_1(t,\tau)\,g(\tau)\,d\tau + \int_0^t\int_0^t K_2(t,\tau_1,\tau_2)\,g(\tau_1)\,g(\tau_2)\,d\tau_1\,d\tau_2 + \cdots \tag{3.15}$$

At $t = 0$, $x(0) = K_0(0) = x_0$. It is clear that $K_0(t)$ is a solution to

$$\frac{dK_0(t)}{dt} + aK_0(t) = b[K_0(t)]^2, \quad K_0(0) = x_0 \tag{3.16}$$

In what follows, we adopt the attitude that $K_0(t)$ is already determined. This can be done by the methods of the preceding chapter, or by any other means.

We put (3.15) into (3.14). This is to hold for all $g(t)$ in some large class, so we equate kernels of like homogeneity across the equation. In this way we obtain

$$\int_0^t K_1(t,\tau)\, g(\tau)\, d\tau = \int_0^t e^{-a(t-\tau)} g(\tau)\, d\tau$$

$$+ 2\int_0^t \int_0^{t'} e^{-a(t-t')} b K_0(t') K_1(t',\tau) g(\tau)\, d\tau\, dt',$$

$$K_1(t,\tau) = 0 \quad \text{for} \quad \tau > t.$$

This is satisfied by setting

$$K_1(t,\tau) = e^{-a(t-\tau)} + 2b\int_0^t e^{-a(t-t')} K_0(t') K_1(t',\tau)\, dt' \tag{3.171}$$

τ acts here only as a parameter.

(3.171) can be solved analytically over any finite interval (Volterra, 1959; also Kolmogorov and Fomin, 1957). The method is simply successive approximations which are shown to converge uniformly over the time interval.

In the same way, we obtain next

$$K_2(t,\tau_1,\tau_2) = \int_0^t e^{-a(t-t')} b\, K_1(t',\tau_1)\, K_1(t',\tau_2)\, dt' \tag{3.172}$$

$$+ 2b \int_0^t e^{-a(t-t')} K_0(t')\, K_2(t',\tau_1,\tau_2)\, dt' \quad ,$$

$$K_2(t,\tau_1,\tau_2) = 0 \text{ for } \tau_1 > t \text{ or } \tau_2 > t.$$

τ_1 and τ_2 act as parameters. We have again the same sort of linear integral equation for K_2 as for K_1. Notice that when $K_0(t) = 0$, we obtain K_1 and K_2 directly. $K_0(t) = 0$ when $x(0) = 0$.

The expression for K_3 is no more involved.

$$K_3(t,\tau_1,\tau_2,\tau_3) = 2b \int_0^t e^{-a(t-t')} K_1(t',\tau_1) K_2(t',\tau_2,\tau_3)\, dt' \tag{3.173}$$

$$+ 2b \int_0^t e^{-a(t-t')} K_0(t')\, K_3(t',\tau_1,\tau_2,\tau_3)\, dt' \quad ,$$

$$K_3(t,\tau_1,\tau_2,\tau_3) = 0 \text{ for } \tau_1, \tau_2 \text{ or } \tau_3 > t.$$

We will not do any more of this, but rather go on to the more general problem covered by equation (3.11).

The equivalent integral equation formulation for (3.11) is

$$x(t) = e^{At} x_0 + \int_0^t e^{A(t-t')} \left\{ f[x(t'),t'] + g(t') \right\} dt' \tag{3.18}$$

We use (3.18), but first set $e^{A(t-t')} = B(t-t')$

to obtain on a component basis

$$x_i(t) = B_{ij}(t)x_{oj} + \int_o^t B_{ij}(t-t')\,g_j(t') + \int_o^t B_{ij}(t-t')\,f^j_{k_1k_2}(t')x_{k_1}(t')x_{k_2}(t')\,dt' + \cdots + \int_o^t B_{ij}(t-t')\,f^j_{k_1\cdots k_n}(t')\,x_{k_1}(t')\cdots x_{k_n}(t')\,dt' \tag{3.191}$$

where it is understood that a repeated index is summed.

We now form a functional representation scheme for $x_i(t)$.

The repeated index summation convention continues in effect.

$$x_i(t) = {}^iK_o(t) + \int_o^t {}^iK_1^k(t,\tau)\,g_k(\tau)\,d\tau + \int_o^t\int_o^t {}^iK_2^{k_1k_2}(t,\tau_1,\tau_2)\,g_{k_1}(\tau_1)\,g_{k_2}(\tau_2)\,d\tau_1\,d\tau_2 + \cdots \tag{3.1201}$$

In the same way as the first example, the introduction of (3.1201) into (3.191) leads to

$$
\begin{aligned}
(3.1211)\quad {}^{i}K_1^{k}(t,\tau) &= B_{ik}(t-\tau) \\
&+ 2\int_0^t B_{ij}(t-t')\, f_{lm}^{j}(t')\, {}^{l}K_0(t')\, {}^{m}K_1^{k}(t',\tau)\, dt' \\
&+ 3\int_0^t B_{ij}(t-t')\, f_{lmn}^{j}(t')\, {}^{l}K_0(t')\, {}^{m}K_0(t')\, {}^{n}K_1^{k}(t',\tau)\, dt' \\
&+ \cdots + \\
&+ n\int_0^t B_{ij}(t-t')\, f_{l_1\cdots l_n}^{j}(t')\, {}^{l_1}K_0(t')\cdots {}^{l_{n-1}}K_0(t')\, {}^{l_n}K_1^{k}(t',\tau)\, dt'
\end{aligned}
$$

where ${}^{i}K_1^{k}(t,\tau) = 0$ for $\tau > t$.

with k and τ fixed and $i = 1, 2, \cdots, n$. k and τ can be considered parameters. (3.1211) is a system of linear integral equations. These can be solved by successive approximations. The theory for this is essentially the same as quoted before (Volterra, 1959).

$K_0(t)$ is the solution to (3.11) when there is no forcing function $g(t)$. We must suppose ${}^{i}K_0(t)$ already determined. Again, if necessary, this can be done through the method indicated by Theorem II. Notice that if $x_i(0) = 0$, ${}^{i}K_0(t) = 0$ and ${}^{i}K_1^{k}(t,\tau)$ can be obtained from (3.1211) directly.

We next determine $\quad {}^{i}K_{2}^{k_1,k_2}(t,\tau_1,\tau_2)$

(3.1212) $\quad {}^{i}K_{2}^{k_1,k_2}(t,\tau_1,\tau_2) =$

$$= \int_0^t B_{ij}(t-t')\, f^{j}_{\ell_1,\ell_2}(t')\, {}^{\ell_1}K_1^{k_1}(t',\tau_1)\, {}^{\ell_2}K_1^{k_2}(t',\tau_2)\, dt'$$

$$+ 3\int_0^t B_{ij}(t-t')\, f^{j}_{\ell_1,\ell_2,\ell_3}(t')\, {}^{\ell_1}K_0(t')\, {}^{\ell_2}K_1^{k_1}(t',\tau_1)\, {}^{\ell_3}K_1^{k_2}(t',\tau_2)\, dt'$$

$$+ \cdots + \frac{n(n-1)}{2!}\int_0^t B_{ij}(t-t')\, f^{j}_{\ell_1\cdots\ell_n}(t') \cdot$$

$$\cdot\, {}^{\ell_1}K_0(t')\cdots{}^{\ell_{n-2}}K_0(t')\, {}^{\ell_{n-1}}K_1^{k_1}(t',\tau_1)\, {}^{\ell_n}K_1^{k_2}(t',\tau_2)\, dt'$$

$$+ 2\int_0^t B_{ij}(t-t')\, f^{j}_{\ell_1,\ell_2}(t')\, {}^{\ell_1}K_0(t')\, {}^{\ell_2}K_2^{k_1,k_2}(t',\tau_1,\tau_2)\, dt'$$

$$+ \cdots + n\int_0^t B_{ij}(t-t')\, f^{j}_{\ell_1\cdots\ell_n}(t')$$

$$\cdot\, {}^{\ell_1}K_0(t')\cdots{}^{\ell_{n-1}}K_0(t')\, {}^{\ell_n}K_2^{k_1,k_2}(t',\tau_1,\tau_2)\, dt'$$

where $\quad {}^{i}K_{2}^{k_1,k_2}(t,\tau_1,\tau_2) = 0 \quad$ for $\quad \tau_1 > t$ or $\tau_2 > t$.

Notice again that if the $x_i(0) = 0$, then ${}^iK_0(t) = 0$, and ${}^iK_2^{k_1,k_2}(t,\tau_1,\tau_2)$ can be determined directly. Also, both ${}^iK_1^k$ and ${}^iK_2^{k_1,k_2}$ are obtained by inverting the same linear operator when ${}^iK_0(t) \neq 0$ for all i. Thus even for this very complicated and general case, one can see that to obtain any order kernel, ${}^iK_n^{k_1,\cdots k_n}$, the same linear operator is to be inverted independent of n.

We will not write out the expression for ${}^iK_3^{k_1,k_2,k_3}$ or the higher order kernels. One proceeds in the same way as above, but of course obtains increasingly more complicated and tedious expressions.

Let us consider ${}^iK_0(t)$ in greater detail. As already mentioned, ${}^iK_0(t)$ is the i^{th} component of the solution when the forcing function n-vector $g(t) = 0$. As such, by the previous chapter, $K_0(t)$ is expandable as

$$(3.122)\quad {}^iK_0(t) = {}^iL_j(t)x_j(0) + {}^iL_{j,j_2}(t)x_j,x_{j_2}(0) + \cdots + {}^iL_{j,\cdots j_n}(t)x_j,(0)\cdots x_{j_n}(0)$$

The expansion converges uniformly over all bounded time and initial conditions. If the problem is non-linear, then correlations down an ensemble involving the solution components will, at least formally, be given in terms of ensemble averages of products of initial position components and forcing function components.

The legality of the formal procedure employed above is given by Theorem III below. In order to state this theorem, we first define a phase space. The space is $2n+1$ dimensional. A point in the phase space is given by the components

$$\left(x_1(0), \cdots, x_n(0), q_1(t), \cdots, q_n(t), t \right)$$

where $q_i \neq h_i$ unless $\int_0^t |q_i(\tau) - h_i(\tau)| d\tau = 0$.

The differential equation (3.11) can be thought of as an operator which assigns to each point of the phase a real number.

A subset $d \times A \times G$ of the phase space is the set of all points such that

$$0 \leq t \leq d$$

$$|x_i(0)| \leq A_i \quad , \quad i = 1, 2, \cdots, n$$

$$|q_i(t)| \leq G_i \quad , \quad i = 1, 2, \cdots, n$$

$$\text{for} \quad 0 \leq t \leq d$$

Now we can state the theorem.

Theorem III

For arbitrary but finite d , A , and G , the representation scheme (3.110i) converges to the solution of (3.11) uniformly over $d \times A \times G$.

The proof is in Appendix C.

(3.2) More About All-Time Representations and Concluding Remarks

In Chapter II, the simple condition that an unforced system damp asymptotically to zero allowed an approximation scheme to converge uniformly over all time. Here, where we are specifically interested in systems with forcing functions, we are unable, at least at this time, to obtain any corresponding result.

The Introduction shows that if an unforced system does not damp asymptotically to but one limit point, then the same system with forcing cannot in general have an all-time Volterra functional representation.

Unfortunately the converse is not true. Consider the system suggested by Professor Edward N. Lorenz.

(3.21)
$$\frac{dx_1}{dt} = -x_1 - x_2^2 + f(t)$$
$$\frac{dx_2}{dt} = -x_2 + x_1 x_2$$

When $f(t)$ is set equal to zero, (3.21) can easily be manipulated into

$$(3.22) \qquad \frac{1}{2}\frac{d}{dt}(x_1^2 + x_2^2) = -(x_1^2 + x_2^2)$$

From this it follows that every solution of the unforced system damps asymptotically to zero.

However, when $f(t) = 2$, the points $(x_1 = 1, x_2 = 1)$ and $(x_1 = 1, x_2 = -1)$ are stable limit points. The point $(x_1 = 2, x_2 = 0)$ is an unstable limit point. The behavior of the system at $t = \infty$ is discontinuous with respect to initial conditions, and hence an all-time representation which would converge uniformly over all time and arbitrarily large but bounded forcing functions and initial conditions is beyond our grasp at the present time.

Nevertheless, it should be emphasized that those cases where statistically stationary forcing functions actually produce asymptotically stationary solution statistics for some definite initial condition, are approachable by the finite time representation scheme of section (3.1) on a computational basis.

Suppose then, that computationally after some finite time, we are able to obtain a reasonably good estimate to the steady statistics

of the solution for a stationary forcing function and a particular initial condition. Now each state of the system within the finite time interval can be considered a new initial condition with the same stationary ensemble of forcing functions. Hence for each point on the finite-time trajectory, and now considered as an initial condition, the stationary statistics are in all cases the same.

The theory for all-time representations of systems with forcing functions is admittedly not yet at a stage at which one is satisfied. However, it is hoped that by machine experiments, the above theory will enable one to mark of those sections of a phase space which under an ergodic ensemble of forcing functions are transformed by the differential equation operator within themselves.

The effort involved in such a project must of necessity be great. Even for problems of the kind considered in Chapter II, where the theory is more developed, any practical problem will require considerable labor. This is an unfortunate, but certainly to be expected, aspect of non-linear analysis. Non-linear problems can present difficulties which do not exist for linear problems. We shall perhaps, simply have to become used to paying more for results in the non-linear case.

Appendix A

The equation first in question is of the form

$$\text{(A1)} \qquad \frac{dx}{dt} = Ax + G_1(t)x + G_2(t)f(x) + g(t)$$

where $x(t)$ and $g(t)$ are n-vectors, $G_1(t)$ and $G_2(t)$ are continuously time dependent nxn matrices, A is a constant nxn matrix, and $f(x)$ is an n-vector whose components are second degree or higher order polynomials in the components of x .

At $t = 0$, $x(0) = x_0$ is the necessary initial condition. (A1) is equivalent to

$$\text{(A2)} \qquad x(t) = e^{At}x_0 + \int_0^t e^{A(t-\tau)}G_1(\tau)x(\tau)\,d\tau + \int_0^t e^{A(t-\tau)}G_2(\tau)f[x(\tau)]\,d\tau + \int_0^t e^{A(t-\tau)}g(\tau)\,d\tau$$

We are supposing that $x(t) = 0$ is an asymptotically stable limit point. Hence the eigenvalues of A all have negative real parts. Every term of $e^{A(t-\tau)}$ is of the form $be^{\lambda(t-\tau)}(t-\tau)^n$ where the real part of λ is negative (Lefschetz, 1957). Each such term is bounded and absolutely integrable.

We use $\| \quad \|$ to denote the norm of a vector or matrix. Unless expressly stated, the norm of a vector or matrix will be taken to be the maximum of the component magnitudes.

Then

$$\lim_{t \to \infty} \int_0^t \| e^{A(t-\tau)} \| \, d\tau$$ exists, is finite,

and we denote it by H .

The nature of $f(x)$ indicates that there is a region about the origin defined by $\| x \| < R$, such that for x in this region, $\| f(x) \| \leq F$ and

$$\| f(x_1) - f(x_2) \| \leq L \| x_1 - x_2 \|$$

where in addition, $F \to 0$ as R^2 and $L \to 0$ with R .

F is the bound for $f(x)$ and L is called a Lipschitz constant.

We next must set up a particular normed space. Let $C^*(b)$ be the set of continuous n-vectors, $\phi^*(t)$, on a closed interval $[0 \leq t \leq d]$. That is, each component of $\phi^*(t)$ is a continuous function of time on $[0, d]$. The norm of $\phi^*(t)$ will be taken as the greatest magnitude of any component for any t in $[0, d]$.

To be more precise,

$$\| \phi^*(t) \| = \sup_i \left\{ \sup_t \left[| \phi_i^*(t) | \right] \right\}$$

We make the additional restriction that for $\phi^*(t)$ to belong to the set $C^*(b)$,

$$\| \phi^*(t) \| \leq b < R$$

With this, $C^*(b)$ is a complete metric space (Kolmogorov and Fomin, 1957).

We wish to show that when $\phi^*(t)$ is substituted into the right side of (A2), we have a contraction mapping. A mapping m of a metric space X into a metric space Y is said to be a contraction mapping if for all x_1 and x_2 in X,

$$\| m(x_1) - m(x_2) \| \leq \alpha \| x_1 - x_2 \|$$

where α is a positive number < 1.

Consider the mapping defined for $[0 \leq t \leq d]$.

(A3)
$$\psi^*(t) = e^{At} x_0 + \int_0^t e^{A(t-\tau)} G_1(\tau) \phi^*(\tau) d\tau + \int_0^t e^{A(t-\tau)} G_2(\tau) f[\phi^*(\tau)] d\tau + \int_0^t e^{A(t-\tau)} g(\tau) d\tau$$

Then $$\left\| {}_a\Psi^*(t) - {}_b\Psi^*(t) \right\| = \sup_i \left\{ \sup_t \left[\left| {}_a\psi_i^*(t) - {}_b\psi_i^*(t) \right| \right] \right\} =$$

$$= \left\| \int_0^t e^{A(t-\tau)} G_1(\tau) \left[{}_a\phi^*(\tau) - {}_b\phi^*(\tau) \right] d\tau \right.$$

$$\left. + \int_0^t e^{A(t-\tau)} G_2(\tau) \left\{ f\left[{}_a\phi^*(\tau) \right] - f\left[{}_b\phi^*(\tau) \right] \right\} \right\|$$

which is $$\leq n^2 H \left[\|G_1\| + \|G_2\| L \right] \left\| {}_a\phi^* - {}_b\phi^* \right\|$$

where $$\|G_\gamma\| = \sup_{t \text{ in } [0,d]} \left[\|G_\gamma(t)\| \right] , \quad \gamma = 1 \text{ and } 2$$

Then for $\|G_1\| < \frac{1}{n^2 H}$ and $\|G_2\|$ simply bounded, there exists an R forcing $L(R)$ to be sufficiently small so that $n^2 H \left[\|G_1\| + L \|G_2\| \right] < 1$. Hence, (A3) is a contraction mapping.

We next show that this mapping takes $C^*(b)$ into itself. From (A3), at the time t',

$$\|\psi^*(t')\| \le n\|e^{At'}\|\,\|x_0\| + n^2\int_0^t \|e^{As}\|\,\|G_1\|\,\|\phi^*\|ds$$
$$+ n^2\int_0^t \|e^{As}\|\,\|G_2\|\,F\,ds$$
$$+ n\int_0^{t'} \|e^{As}\|\,\|g\|\,ds$$

Here $\|g\| = \sup_t [\|g(t)\|]$

As mentioned above, there is a bound B such that

$$\|e^{At'}\| \le B \quad \text{for all } t' \text{ in } [0 \le t' < \infty)$$

Thus $$\|\psi^*(t')\| \le nB\|x_0\| + nH\|g\|$$
$$+ n^2H[\|G_1\|b + \|G_2\|F]$$

Since $F \to 0$ as b^2, for $\|x_0\|$, $\|G_1\|$ and $\|g\|$ small enough, there exists b such that $\|\psi^*(t')\| < b$ for all t'. This shows that the image of the contraction mapping given by (A3) on $C^*(b)$ is in $C^*(b)$ for some b with suitable restrictions on the initial conditions, the matrix $G_1(t)$ and the forcing function $g(t)$.

There is a contraction mapping theorem (Kolmogorov and Fomin, 1957) which ways that if m is a contraction mapping of the complete metric space $C^*(b)$ into itself, then the equation $\phi = m[\phi]$ possesses a unique solution. Further, ϕ can be found by successive approximations. That is if ${}_1\phi$ is any element of $C^*(b)$, and we define

$$ {}_2\phi = m[{}_1\phi] \quad , \quad {}_3\phi = m[{}_2\phi] = m^2[{}_1\phi] \qquad \text{etc.} $$

then $$ \phi = \lim_{n \to \infty} m^n[{}_1\phi] $$

This means that for (A3) we can define

$$ {}_1x(t) = e^{At}x_0 + \int_0^t e^{A(t-\tau)}g(\tau)\,d\tau $$

and recursively obtain

$$ {}_{n+1}x(t) = e^{At}x_0 + \int_0^t e^{A(t-\tau)}G_1(\tau)\,{}_nx(\tau)\,d\tau $$
$$ + \int_0^t e^{A(t-\tau)}G_2(\tau)f[{}_nx(\tau)]\,d\tau + \int_0^t e^{A(t-\tau)}g(\tau)\,d\tau $$

The successive approximations approach $x(t)$, the solution vector, uniformly over $[0, d]$. That is, for any ϵ there exists an $N(\epsilon)$ such that for all $n > N(\epsilon)$

$$\left\| x(t) - {}_n x(t) \right\| < \epsilon \text{ for all } t \text{ in } [0, d] .$$

Finally, although the functions in $C^*(b)$ have been supposed defined only on $[0, d]$, the conditions shown sufficient to insure (A3) as a contraction mapping of $C^*(b)$ into itself have none of them been dependent on d. Thus the successive approximations converge to the solution vector uniformly over all time.

One generalization follows easily. It has been required of $G_2(t)$ only that each component be bounded over time. If equation (A1) is replaced by

$$\text{(A4)} \qquad \frac{dx}{dt} = Ax + G_1(t)x + G_2(t)f_2(x) + G_3(t)f_3(x) + \cdots + G_n(t)f_n(x) + g(t)$$

where each $G_2, G_3, \ldots, G_n$ is bounded uniformly over time, it is clear that the above results are still valid.

The vector sum

$$G_2(t)f_2(x) + \cdots + G_n(t)f_n(x)$$ can be combined into the single vector $H[t, x(t)]$. Each component of $H[t, x]$ will be a finite polynomial of degree 2 or greater

in the components of x , and the coefficients will be uniformly bounded functions of time. With this, (A4) becomes

$$\text{(A5)} \qquad \frac{dx}{dt} = Ax + G(t)x + H[t, x] + g(t)$$

Another generalization can be obtained. Consider the linear equation

$$\text{(A6)} \qquad \frac{dx}{dt} = B(t)x + g(t)$$

If the solution to (A6) is bounded for every bounded initial condition x_0 and every bounded $g(t)$, then the same kind of results as obtained above hold for

$$\text{(A7)} \qquad \frac{dx}{dt} = B(t)x + H[t, x] + g(t)$$

We will not go through a proof, but simply state a result which allows one to use the contraction mapping method as before.

(A7) is equivalent to

$$\text{(A8)} \qquad x(t) = X(t)x_0 + X(t)\int_0^t X^{-1}(\tau) H[\tau, x(\tau)]\, d\tau + X(t)\int_0^t X^{-1}(\tau) g(\tau)\, d\tau$$

where $X(t)$ is the solution matrix for (A6) with $g(t)$ set equal to zero.

Under the condition that (A6) give bounded solutions for bounded forcing functions, Bellman (1948) has shown that

$$\int_0^t \| X(t) X^{-1}(\tau) \| d\tau \quad \text{is bounded for all } t \text{ ,}$$

and that $\lim_{t \to \infty} \| X(t) \| = 0$. This last implies of course that $\| X(t) \|$ is also bounded for all time since $X(t)$ is continuous.

The contraction mapping method goes through as before, and we are now able to summarize our results in the following theorem.

Theorem

If for equation (A6)

$$\text{(A6)} \qquad \frac{dx}{dt} = B(t)x + g(t)$$

every bounded forcing function $g(t)$, with bounded initial conditions x_0 , produces a bounded solution, then there exists an $R > 0$ and a $g > 0$ such that for $\| x_0 \| < R$ and $\| g(t) \| < g$ for all t , the equation (A7)

(A7) $$\frac{dx}{dt} = B(t)x + H[t, x] + g(t)$$

can be solved by successive approximations. The approximations are determined by the linear equation

(A9) $$\frac{d}{dt}({}_{n+1}x) = B(t)\,{}_{n+1}x + H[t, {}_{n}x] + g(t)$$

with ${}_{n+1}x(0) = x_0$; and ${}_{n}x(t)$

converges uniformly over all time to the solution $x(t)$.

A special but important case arises when $B(t)$ is a constant matrix plus a pertubation time dependent matrix.

If the eigenvalues of A all have negative real parts, there exists a $G > 0$ such that for $\|G(t)\| < G$ every solution of

(A10) $$\frac{dx}{dt} = Ax + G(t)x + g(t)$$

is bounded for bounded $g(t)$ and x_0 ; and hence the hypothesis of the theorem is satisfied.

Appendix B

Asymptotically Stable Systems

We will deal with a system

(B1) $$\frac{dx}{dt} = f(x,t)$$

where x and f are n-vectors and $f(0,t) = 0$. Further, it is supposed that the solution $x(t)$ to (B1) is asymptotically stable.

Let R be an arbitrary but fixed bound on the initial conditions. That is

$$\|x(0)\| = \max_i \left[|x_i(0)|\right] \leq R$$

Then the set $S(R)$ of initial conditions defined by $\|x(0)\| \leq R$ is compact.

Let p be any point in $S(R)$. By supposition, there exists a $T(p, \delta/2)$ such that the trajectory originating at p at $t = 0$ is within $\delta/2$ of zero for all $t > T(p, \delta/2)$. By known results (Lefschetz, 1957), there exists an open neighborhood $N(p, \delta/2)$ with the property that all points q in $N(p, \delta/2)$

at $t = 0$ are within $\delta/2$, at $t = T$, of the trajectory originating at p . Hence all q in $N(p, \delta/2)$ are within δ of zero at $t = T$.

For each point p in $S(R)$ there exists such a neighborhood $N(p, \delta/2)$. Since $S(R)$ is compact, a finite number of these neighborhoods cover $S(R)$. It follows that for arbitrary $\delta > 0$ and $R > 0$, there exists a $t_1(\delta, R)$ such that all trajectories $x(t)$, with $\| x(c) \| \leq R$, end up uniformly $\| x(t) \| < \delta$ for all $t > t_1$.

Appendix C

The fundamental equation with which we deal is

$$\frac{dx}{dt} = f(x,t) + g(t) \tag{C1}$$

where x, f, and g are n-vectors. $f(x,t)$ is analytic in $(x_1, \ldots, x_n, t)$. However, we will assume for convenience that the form of the i^{th} component of f is

$$f_i(x,t) = f^i_j(t)\, x_j + f^i_{j_1 j_2}(t) x_{j_1} x_{j_2} + \cdots + f^i_{j_1 \cdots j_n}(t) x_{j_1} \cdots x_{j_n} \tag{C2}$$

For (C2), and the rest of Appendix C, the summation convention applies.

For bounded time, the $f^i_j(t), \ldots, f^i_{j_1 \cdots j_n}(t)$ are bounded functions of time.

The problem is defined over the finite time interval $[0 \le t \le d]$. It is assumed that over $[0, d]$, bounded forcing functions produce bounded solutions to (C1). Let G be the bound for $g(t)$, and

R the bound for $x(t)$. That is,

$$\left[|x_i(t)|\right] \leq R \qquad \text{for all } \left[c \leq t \leq d\right] \quad i = 1, \ldots, n$$

when $$\left[|g_i(t)|\right] \leq G \qquad \text{for all } \left[0 \leq t \leq d\right] \quad i = 1, \ldots, n$$

There is then a bound F for $f(x, t)$.

$$\left[|f_i(x,t)|\right] \leq F \qquad \text{for all } \left[0 \leq t \leq d\right] \quad i = 1, \ldots, n$$

$$\text{and all } |g_j(t)| \leq G \quad j = 1, \ldots, n$$

Further, there exists a Lipschitz constant L such that

$$|f_i(x,t) - f_i(y,t)| < L \max_j |x_j(t) - y_j(t)|$$

at any t in $[0, d]$ and all x, y bounded by R.

Now (C1) is equivalent to

$$\text{(C3)} \qquad x(t) = x_0 + \int_0^t f[x(\tau), \tau]\, d\tau + \int_0^t g(\tau)\, d\tau$$

We wish to show that the right side of (C3) defines an operator which is continuous on a certain topological space to be defined later.

Consider first the quantity $[x(t) - y(t_1)]$, where $y(t_1)$ is defined by

$$(C4) \qquad y(t_1) = y_0 + \int_0^{t_1} f[y(\tau), \tau]\, d\tau + \int_0^{t_1} h(\tau)\, d\tau$$

Let, for definiteness, $t_1 \geq t$. Then with the above definitions of bounds, the following inequality will hold.

$$(C5) \qquad |x_i(t) - y_i(t_1)| \leq |x_i(0) - y_i(0)| + \left| \int_0^{t_1} [g_i(\tau) - h_i(\tau)]\, d\tau \right| + L \max_{j,t} \left[|y_j(t) - x_j(t)| \right] t_1 + (F+G)(t - t_1)$$

(C5) suggests a topology for the space of forcing functions. We say that $h(t)$ is in an ϵ neighborhood of $g(t)$ if and only if

$$\left| \int_a^b [g_i(\tau) - h_i(\tau)]\, d\tau \right| < \epsilon$$

for all i and any $[a, b]$ in the fundamental interval $[0, d]$. Actually, this is a well known topology, and is called the weak* topology induced on $L'[0, d]$. It is also a known result that the set of all $L'[0, d]$ integrable functions, $g(t)$, for which $|g_i(t)| \leq G$ for all t in $[0, d]$ and all i ,

is a compact set under the weak* topology defined above. (Dumford and Schwartz, 1958; corollary IV 8.11).

We return to (C5) and restrict t to be in $[C, \lambda]$ where $\lambda L = \alpha < 1$. Notice that since $y_i(t_1)$ is continuous on $[0, d]$, it is uniformly continuous there. Now let an arbitrary $\epsilon_1 > C$ be chosen. Then there exists $\delta > 0$ such that $|y_i(t) - y_i(t_1)| < \epsilon_1$ for all i whenever $|t - t_1| < \delta$. Hence for $|t - t_1| < \delta$ and all t in $[0, d]$, and with $|x_i(0) - y_i(0)| < \epsilon_2$ for all i, and $\left| \int_0^{t_1} [g_i(\tau) - h_i(\tau)\, d\tau \right| < \epsilon_3$ for all i, and all t_1 in $[0, t]$,

then $\max_i |x_i(t) - y_i(t_1)| < \dfrac{\epsilon_1 + \epsilon_3 + \alpha\epsilon_2 + (F + G)\delta}{1 - \alpha}$

which is $< \epsilon$ for ϵ_1, ϵ_2, ϵ_3 and δ chosen small enough. λ, δ, ϵ_1, ϵ_2 and ϵ_3, do not depend on t.

We can break $[C, d]$ into a finite number of overlapping intervals each of length no greater than λ. The immediately preceding analysis can be produced on each sub-interval. All this implies that

for arbitrary $\epsilon > 0$, there exists a $\delta(\epsilon)$ such that

$|x_i(t) - y_i(t_1)| < \epsilon$ for all i at any t in $[0, d]$ whenever

(C6a,b,c) $|x_i(0) - y_i(0)| < \delta$ for all i

$$|t - t_1| < \delta$$

and $$\left|\int_0^t [g_i(T) - h_i(T)]\,dT\right| < \delta.$$

for all i and all t in $[0, d]$.

The product space S defined as the set of (2n+1)-vectors

$$(t, x_1(0), \cdots, x_n(0), g_1, \cdots, g_n)$$

is given the topology indicated by (C6a,b,c). This is a product topology on a product space. Hence the set $d \times A \times G$ of points in S such that

$$0 \le t \le d$$

$$|x_i(0)| \le A \quad \text{for all } i$$

$$|g_i(t)| \le G \quad \text{for all } i \text{ and all } t \text{ in } [0, d]$$

is a compact set under the product topology defined.

It is an immediate consequence of Brilliant's (1958) results that the Volterra functional polynomials

$$K_0[t, x_0] + \int_0^d K_i^1[t, x_0, \tau]\, g_i(\tau)\, d\tau$$

$$+ \cdots + \int_0^d \cdots \int_0^d K_{i_1 \cdots i_n}^n [t, x_0, \tau_1, \ldots, \tau_n] \cdot$$

$$\cdot\, g_{i_1}(\tau_1) \cdots g_{i_n}(\tau_n)\, d\tau_1 \cdots d\tau_n$$

are continuous on S with respect to the product topology defined through (C6a,b,c). Since $d \times A \times G$ is compact, and since it is clear that the functional polynomials separate points of S, by invoking the Stone-Weierstrass theorem, we can conclude that the solution to (C1) can be approximated uniformly over $d \times A \times G$ by a finite system of Volterra functional polynomials.

ACKNOWLEDGEMENT

This work is an outgrowth of conversations held with Professor Martin Schetzen of the Department of Electrical Engineering, M.I.T., over the summer of 1958. The author is indebted to Professor Schetzen for the stimulation of these discussions, as well as for his continuing interest.

Professor Edward N. Lorenz of the Department of Meteorology, M.I.T., acted as supervisor. It is an honor to acknowledge his support and sympathetic attitude toward basic research. Without the truly academic atmosphere pervading the Statistical Forecasting Project, of which Professor Lorenz is Director, an undertaking of this kind would not have been possible.

The author wishes to thank Professor Stephen Simpson of the Department of Geology and Geophysics, M.I.T., and Professor Victor Starr of the Department of Meteorology, M.I.T., for their time and personal interest.

Miss Marie Louise Guillot typed the text from most illegible manuscript. The author wishes to thank her for her considerable pains. Also, the author expresses his gratitude to Miss Ellen Fetter who assisted in the preparation of the report.

References

Apostol, T.M., 1957, Mathematical Analysis, Addison-Wesley, Reading, Mass.

Batchelor, G.K., 1960, The Theory of Homogeneous Turbulence, Cambridge University Press, Cambridge, Great Britain

Bellman, R., 1948, On an Application of a Banach-Steinhaus Theorem to the Study of the Boundedness of Solutions of Nonlinear Differential and Difference Equations, Ann. of Math., 49, 515-522.

Birdgland, T.T., 1961, Asymptotic Behavior of the Solutions of Nonhomogeneous Differential Equations, Proc. Amer. Math. Soc., 12, 546-552.

Bochner, S. and Martin, W.T., 1948, Several Complex Variables, Princeton University Press.

Brilliant, M.B., 1958, Theory of the Analysis of Nonlinear Systems, M.I.T. Res. Lab. of Elec. Technical Report 345.

Coddington, E.A. and Levinson, N., 1955, Theory of Ordinary Differential Equations, McGraw-Hill, N. Y.

Dunford, N. and Schwartz, J.T., 1958, Linear Operators, Interscience Publishers, N. Y.

Friedman, B., 1956, Principles and Techniques of Applied Mathematics, John Wiley and Sons, N. Y.

Jeffreys, H., 1931, Cartesian Tensors, Cambridge University Press, Cambridge, Great Britain.

Kolmogorov, A.N. and Fomin, S.V., 1957, Elements of the Theory of Functions and Functional Analysis, Graylock Press, Rochester, N. Y.

Lamb, H., 1945, Hydrodynamics, Dover Publications, N. Y.

Lefschetz, S., 1957, Differential Equations: Geometric Theory, Interscience Publishers, N. Y.

Lighthill, M.J., 1959, An Introduction to Fourier Analysis and Generalized Functions, Cambridge University Press, Cambridge, Great Britain.

Lorenz, E.N., 1953, The Interaction Between a Main Flow and Random Disturbances, Tellus, 6, 238-250.

Minorsky, N., 1947, Introduction to Non-Linear Mechanics, J.W. Edwards, Ann Arbor, Mich.

Schetzen, M., 1962, Some Problems in Nonlinear Theory, M.I.T. Res. Lab. of Elec. Report 390.

Volterra, V., 1959, Theory of Functionals and of Integral and Integro-Differential Equations, Dover, N. Y.

Wehausen, J.V. and Laitone, E.V., 1960, Surface Waves, Handbuch der Physik, 9, 446-778.

Wiener, N., 1958, Nonlinear Problems in Random Theory, M.I.T. Technology Press, Cambridge, Mass. and John Wiley and Sons, N. Y.

Statistical Forecasting Project, Massachusetts Institute of Technology
Cambridge, Massachusetts

Report No. AFCRL-62-1137

REPRESENTATION SCHEMES FOR INVESTIGATING NON-LINEAR PROCESSES

Donald B. DeVorkin

Scientific Report No. 1, December 1962, 92 pages, Air Force Contract No. AF19(604)-4969, Unclassified Report.

The work is concerned with systems of ordinary differential equations. A framework is developed in which a statistical approach to the analysis of the equations is natural. One seeks the correlations of the solution in terms of the correlations of initial conditions and/or forcing functions. General representation schemes are developed for this purpose. It is shown that general schemes can converge for arbitrarily large but finite time. In some special cases, all-time representations are obtained.

Unclassified

1. Differential Equations with Random Initial Conditions and Random Forcing
2. General Solutions of Non-Linear Differential Equations
3. DeVorkin, Donald B.

Statistical Forecasting Project, Massachusetts Institute of Technology
Cambridge, Massachusetts

Report No. AFCRL-62-1137

REPRESENTATION SCHEMES FOR INVESTIGATING NON-LINEAR PROCESSES

Donald B. DeVorkin

Scientific Report No. 1, December 1962, 92 pages, Air Force Contract No. AF19(604)-4969, Unclassified Report.

The work is concerned with systems of ordinary differential equations. A framework is developed in which a statistical approach to the analysis of the equations is natural. One seeks the correlations of the solution in terms of the correlations of initial conditions and/or forcing functions. General representation schemes are developed for this purpose. It is shown that general schemes can converge for arbitrarily large but finite time. In some special cases, all-time representations are obtained.

Unclassified

1. Differential Equations with Random Initial Conditions and Random Forcing
2. General Solutions of Non-Linear Differential Equations
3. DeVorkin, Donald B.

ECTF

Statistical Forecasting Project, Massachusetts Institute of Technology
Cambridge, Massachusetts

Report No. AFCRL-62-1137

REPRESENTATION SCHEMES FOR INVESTIGATING NON-LINEAR PROCESSES

Donald B. DeVorkin

Scientific Report No. 1, December 1962, 92 pages, Air Force Contract No. AF19(604)-4969, Unclassified Report.

The work is concerned with systems of ordinary differential equations. A framework is developed in which a statistical approach to the analysis of the equations is natural. One seeks the correlations of the solution in terms of the correlations of initial conditions and/or forcing functions. General representation schemes are developed for this purpose. It is shown that general schemes can converge for arbitrarily large but finite time. In some special cases, all-time representations are obtained.

Unclassified

1. Differential Equations with Random Initial Conditions and Random Forcing
2. General Solutions of Non-Linear Differential Equations
3. DeVorkin, Donald B.

Statistical Forecasting Project, Massachusetts Institute of Technology
Cambridge, Massachusetts

Report No. AFCRL-62-1137

REPRESENTATION SCHEMES FOR INVESTIGATING NON-LINEAR PROCESSES

Donald B. DeVorkin

Scientific Report No. 1, December 1962, 92 pages, Air Force Contract No. AF19(604)-4969, Unclassified Report.

The work is concerned with systems of ordinary differential equations. A framework is developed in which a statistical approach to the analysis of the equations is natural. One seeks the correlations of the solution in terms of the correlations of initial conditions and/or forcing functions. General representation schemes are developed for this purpose. It is shown that general schemes can converge for arbitrarily large but finite time. In some special cases, all-time representations are obtained.

Unclassified

1. Differential Equations with Random Initial Conditions and Random Forcing
2. General Solutions of Non-Linear Differential Equations
3. DeVorkin, Donald B.

Statistical Forecasting Project, Massachusetts Institute of Technology
Cambridge, Massachusetts

Report No. AFCRL-63-1137

REPRESENTATION SCHEMES FOR INVESTIGATING NON-LINEAR PROCESSES
Donald B. DeVorkin

Scientific Report No. 1, December 1962, 92 pages, Air Force Contract No. AF19(604)-4969, Unclassified Report.

The work is concerned with systems of ordinary differential equations. A framework is developed in which a statistical approach to the analysis of the equations is natural. One seeks the correlations of the solution in terms of the correlations of initial conditions and/or forcing functions. General representation schemes are developed for this purpose. It is shown that general schemes can converge for arbitrarily large but finite time. In some special cases, all-time representations are obtained.

Unclassified

1. Differential Equations with Random Initial Conditions and Random Forcing
2. General Solutions of Non-Linear Differential Equations
3. DeVorkin, Donald B.

Statistical Forecasting Project, Massachusetts Institute of Technology
Cambridge, Massachusetts

Report No. AFCRL-63-1137

REPRESENTATION SCHEMES FOR INVESTIGATING NON-LINEAR PROCESSES
Donald B. DeVorkin

Scientific Report No. 1, December 1962, 92 pages, Air Force Contract No. AF19(604)-4969, Unclassified Report.

The work is concerned with systems of ordinary differential equations. A framework is developed in which a statistical approach to the analysis of the equations is natural. One seeks the correlations of the solution in terms of the correlations of initial conditions and/or forcing functions. General representation schemes are developed for this purpose. It is shown that general schemes can converge for arbitrarily large but finite time. In some special cases, all-time representations are obtained.

Unclassified

1. Differential Equations with Random Initial Conditions and Random Forcing
2. General Solutions of Non-Linear Differential Equations
3. DeVorkin, Donald B.

Statistical Forecasting Project, Massachusetts Institute of Technology
Cambridge, Massachusetts

Report No. AFCRL-63-1137

REPRESENTATION SCHEMES FOR INVESTIGATING NON-LINEAR PROCESSES
Donald B. DeVorkin

Scientific Report No. 1, December 1962, 92 pages, Air Force Contract No. AF19(604)-4969, Unclassified Report.

The work is concerned with systems of ordinary differential equations. A framework is developed in which a statistical approach to the analysis of the equations is natural. One seeks the correlations of the solution in terms of the correlations of initial conditions and/or forcing functions. General representation schemes are developed for this purpose. It is shown that general schemes can converge for arbitrarily large but finite time. In some special cases, all-time representations are obtained.

Unclassified

1. Differential Equations with Random Initial Conditions and Random Forcing
2. General Solutions of Non-Linear Differential Equations
3. DeVorkin, Donald B.

Statistical Forecasting Project, Massachusetts Institute of Technology
Cambridge, Massachusetts

Report No. AFCRL-63-1137

REPRESENTATION SCHEMES FOR INVESTIGATING NON-LINEAR PROCESSES
Donald B. DeVorkin

Scientific Report No. 1, December 1962, 92 pages, Air Force Contract No. AF19(604)-4969, Unclassified Report.

The work is concerned with systems of ordinary differential equations. A framework is developed in which a statistical approach to the analysis of the equations is natural. One seeks the correlations of the solution in terms of the correlations of initial conditions and/or forcing functions. General representation schemes are developed for this purpose. It is shown that general schemes can converge for arbitrarily large but finite time. In some special cases, all-time representations are obtained.

Unclassified

1. Differential Equations with Random Initial Conditions and Random Forcing
2. General Solutions of Non-Linear Differential Equations
3. DeVorkin, Donald B.